

Министерство науки и высшего образования
Российской Федерации
Алтайский государственный университет
Институт математики и информационных технологий

С.В. Дронов

**ТЕОРИЯ И СТАТИСТИКА СЛУЧАЙНЫХ
ПРОЦЕССОВ**
в задачах прикладного анализа

Учебное пособие

Барнаул, 2022

Об издании 1, 2

Сведения об издании

УДК 519.21
ББК 22.171.5
Д 758

Автор: **Сергей Вадимович Дронов**, к.ф. - м.н.

Рецензент: Григорий Владимирович Пышнограй, д.ф.-
м.н., профессор АлтГТУ

Д 758 Дронов, С.В. Теория и статистика случайных процессов в задачах прикладного анализа: учебное пособие / С.В. Дронов; Алтайский государственный университет. – Барнаул: АлтГУ, 2022. – 1 CD-R (1 Мб). – Систем.требования: IntelPentium, 1,6 GHz и более; 512 Мб (RAM) и более; Microsoft Windows 7 и выше; AdobeReader. – Загл. с титул. экрана. – Текст: электронный.

Учебное электронное издание

Предлагаемое учебное пособие содержит краткое изложение основных положений теории случайных процессов. Особый акцент сделан на приложениях приводимых результатов, хотя для большинства из них дается строгое обоснование. Пособие основано на учебных академических курсах «Теория случайных процессов» и «Теория случайных процессов в задачах прикладного анализа», которые автор более 10 лет читает для магистрантов института математики и информационных технологий АлтГУ.

УДК 517.0
ББК 22.172

© С.В. Дронов, 2022.

© Алтайский государственный университет, 2022.

производственно-технические сведения

Редактор: О.В. Котова
Верстка: О.А. Жданова

Дата подписания к использованию: 06.04.2022

Объем издания: 1 Мб
Комплектация издания: 1 CD-R

Тираж 15 дисков

ФГБОУ ВО «Алтайский государственный университет»
656049, Барнаул, пр. Ленина, 61

Содержание

1. Случайный процесс – что это?
2. Основные математические структуры теории случайных процессов
3. Цели курса и элементы его траектории
4. Гауссовские случайные процессы
5. Некоторые важные классы случайных процессов
6. Понятие марковского процесса
 - 6.1 Концепция марковости
 - 6.2 Общая модель эволюции цепи Маркова
 - 6.3 Эргодические цепи Маркова
7. Линейные преобразования случайных процессов
 - 7.1 Гильбертово пространство случайных величин
 - 7.2 Дифференцирование и интегрирование в среднем квадратическом
 - 7.3 Стохастический интеграл неслучайной функции
8. Спектральная теория
 - 8.1 Спектральная функция
 - 8.2 Формула Котельникова – Шеннона
 - 8.3 Интегралы, производные
9. Прогнозирование в частных случаях
 - 9.1 Постановка задачи прогноза
 - 9.2 Сингулярный случай
 - 9.3 Прогноз в регулярном случае
 - 9.3 Оценка спектральной плотности
10. Процессы размножения и гибели
 - 10.1 Основные предположения
 - 10.2 Одномерные распределения
 - 10.3 Примеры
11. О непрерывных модификациях
12. Условные математические ожидания
13. Мартингалы
 - 13.1 Определения
 - 13.2 Полнота пространства мартингалов
 - 13.3 Стохастический интеграл по мартингалу

- 14. **Интегралы Ито**
 - 14.1 **Стохастические интегралы Ито**
 - 14.2 **Стохастический дифференциал**
 - 15. **Задачи для самостоятельного решения**
- Библиографический список**

Глава 1

Случайный процесс – что это?

Перед вами – очередной учебник по очередному вероятностному курсу. Традиционная схема высшего математического образования подразумевает продвижение по системе изучаемых понятий от простого к сложному. Цепочка наших вероятностных курсов практически ни на шаг не отступает от этой схемы. Поясню: понятие случайного события (которое сегодня, между прочим, начинает изучаться в 5 классе средней школы) можно назвать простейшей бинарной ситуацией. Действительно, событие может произойти, а может и нет. Другие варианты в классической теории вероятностей исключаются, хотя сегодня уже имеется достаточно много строгих математических теорий, которые изучают нечто промежуточное (например, [1, 2, 3]). Далее на сцену выходят случайные величины и случайные векторы, изучение которых приводит к необходимости ввести в сферу наших интересов все действительные числа, а затем и многомерные пространства произвольной конечной размерности. Эти множества выступают в качестве набора всех возможных результатов эксперимента, в котором наблюдаются значения, принимаемые исследуемой случайной величиной или вектором.

Но и этого набора «основных пространств» для описания возможных результатов эксперимента становится недостаточно, когда мы переходим к изучению курса математической статистики. Дело в том, что эта наука пытается дать универсальные рецепты принятия решений по выборкам произвольного, чаще всего, до-

вольно большого объема. Необходимость рассматривать результаты нескольких экспериментов и потребность сравнивать их результаты приводит к идее R^∞ , пространства бесконечного количества измерений, в которое могут быть помещены выборки любых конечных объемов. Правда, для соблюдения формальных моментов, каждая такая выборка должна превращаться не в одну точку этого пространства, а в некоторое алгебраическое многообразие в нем, имеющее конечную коразмерность. Это происходит за счет того, что только небольшое количество непосредственно наблюдаемых координат бесконечномерного вектора, в который превращается наша конкретная выборка, фиксированы. Остальные могут выбираться произвольным образом, – и мы получаем копию пространства R^∞ «сдвинутую» относительно нуля на конечномерный вектор.

Отметим также, что, несмотря на обозначение, размерность вводимого для нужд математической статистики R^∞ является минимальной среди всех бесконечномерных пространств. Оно, если можно так выразиться, счетномерное. Осознав всю неочевидность возникшей системы математических конструкций, теперь нетрудно понять, почему ученые, работающие в этой области, все чаще отказываются называться статистиками, а хотят быть названными специалистами в области интеллектуального анализа данных.

К понятию случайного процесса приводит желание наблюдать процесс изменения распределения случайной величины с течением времени. Естественно, нельзя обойтись без произведения ряда экспериментов по ее наблюдению. Но при этом мы должны понимать, что каждый раз мы наблюдаем, вообще говоря, новую случайную величину, – ведь наблюдения производятся в разное время. Конечно, это то же самое, что наблюдать несчетное семейство случайных величин. Каждому значению времени может соответствовать свое распределение. Поэтому даже при решении простейших задач такого типа мы вынуждены рассматривать наблюдения, результат которых расположен уже в пространстве с несчетным, континуальным числом измерений. Более того, мы просто не в состоянии получить более одного результата наблюдений каждого из возможных в изучаемой эволюции распределений, ведь время не остановишь. Изучениями таких вот неочевидных вещей мы и займемся. И, хотя имеется огромное количество теоретических руководств по очерченному кругу вопросов (см. библиографию), надо понимать, что практический аспект в них, как и в задачах математической статистики, может появиться только на основе числовых наблюдений.

Та часть теории, в которой приводятся конкретные рекомендации к решению практических задач, называется статистикой случайных процессов.

Примеров случайных процессов в окружающей нас реальности немало. Например, хотя напряжение в городской электрической сети равно 220 вольт, но на самом деле в каждый конкретный момент времени оно меняется под влиянием целого ряда непредсказуемых факторов. Более того, всем известны случаи скачков напряжения в сети, достаточно больших по величине, приводящих в результате к перегораниям включенных в сеть электроприборов. Таким образом, напряжение в сети в момент времени t есть пример реального случайного процесса.

Ниже мы столкнемся с диффузионными процессами, которые описывают, скажем, процент определенного вещества, успевшего проникнуть к данному моменту через пористую перегородку. Еще одним примером, который также рассматривается в нашем курсе, служат так называемые процессы размножения и гибели, где изучается количество некоторых организмов или частиц, находящихся в данный момент времени в некоторой ограниченной области.

В последнем из примеров можно обратить внимание на присутствующий в нем определенный момент дискретности. Например, подсчет количества частиц в области, как правило, производится не непрерывно, а с некоторой периодичностью. Если это действительно так, то этот процесс, видимо, удобнее изучать как так называемый временной ряд, теория которых разработана существенно подробнее, чем общая теория случайных процессов. Основным отличием теории временных рядов от теории случайных последовательностей, например, служит именно то, что во временных рядах предполагается наличие некоторого неизменного шага, промежутка между последовательными измерениями.

Глава 2

Основные математические структуры теории случайных процессов

Пусть T – множество произвольной природы. Через R^T принято обозначать пространство всех действительныхзначных функций, определенных на T . Подобное немного странное обозначение проще всего, вероятно, объяснить так. Если, скажем, T – множество из 3 элементов x, y, z , то любую функцию, заданную на T и принимающую действительные значения, можно задать с помощью таблицы 2 на 3: первая строка содержит список этих элементов, а во второй, под каждым из них указано то значение, которое принимает определяемая функция. Поскольку для различных функций меняется только вторая строка, то между всеми возможными функциями на T и всеми такими строками существует биекция. Осталось отметить, что набор всевозможных строк длины 3 – это и есть трехмерное евклидово пространство R^3 . Если же T является произвольным множеством, например, отрезком действительной прямой, то процесс задания функций на T тоже можно представить, как подписывание некоторого числа под каждым из его элементов (хотя физически это, разумеется, невозможно). В результате каждая из действительныхзначных функций может быть отождествлена со

строкой из действительных чисел «той же длины», что и T . А набор всех таких строк уже может быть обозначен R^T – по аналогии с конечным случаем.

Нужно здесь объяснить еще один момент, который будет полезен далее: почему для множества T , например, имеющего вид $\{x, y, z\}$ мы фактически пришли к обозначению $R^{\{x,y,z\}}$, а не применили стандартное R^3 ? Для того, чтобы в этом разобраться, давайте предположим, что в одной задаче с $T = \{1, 3, 8\}$ нам нужно также рассмотреть $U = \{1, 3, 8, 9\}$ и $V = \{1, 3, 5, 8\}$. Оба этих множества содержат множество T как подмножество и оба состоят из 4-х элементов. Но ясно, что $R^U \neq R^V$, хотя R^T очевидно, может быть естественным образом вложено в оба этих четырехмерных пространства. Чтобы трехмерный вектор $(x, y, z) \in R^T$ рассматривать, как элемент R^U , нужно, например, добавить в конце искусственную координату: $(x, y, z, 0) \in R^U$. Но вот, чтобы поместить этот же вектор в R^V искусственную координату нужно поместить в другом месте: $(x, y, 0, z)$. Таким образом, чтобы совместить все эти конструкции в одной задаче, удобно будет все наши пространства рассматривать как подпространства R^W , где $W = \{1, 3, 5, 8, 9\}$ – объемлющее множество из 5 элементов. При этом элементы R^U будут в этом пространстве содержаться в виде $(\cdot, \cdot, 0, \cdot, \cdot)$, а элементы R^V – в виде $(\cdot, \cdot, \cdot, \cdot, 0)$. Разумеется, элементам R^T соответствует представление $(\cdot, \cdot, 0, \cdot, 0)$.

Рассмотрим далее вероятностное пространство $\langle \Omega, \mathcal{F}, \mathbf{P} \rangle$. Отображение $\xi : T \times \Omega \rightarrow R$ назовем *случайной функцией*, если при фиксировании любого из $t \in T$ оно становится случайной величиной. Иначе это записывают так: $\forall t \in T \quad \xi(t) = \xi(t, \cdot)$ – случайная величина. Как всегда, случайный аргумент в левой части формулы опущен.

Точка вместо одного из аргументов обычно означает, что остальные аргументы фиксируются, а то, что остается, мы рассматриваем как условное обозначение функции замененного точкой аргумента. Ясно также, что $\xi(\cdot, \omega) \in R^T$. Иногда требуется ограничить семейство функций, которые допустимы при фиксировании случайного аргумента, – это мы каждый раз будем оговаривать особо.

Если $T \subset R$ и параметр $t \in T$ интерпретируется как время, то случайную функцию называют *случайным процессом*. Когда мы сталкиваемся с ситуацией, где T представляет собой множество целых чисел \mathbf{Z} или натуральных чисел \mathbf{N} , то мы называем случайный процесс *случайной последовательностью*. Это, разумеется, то же самое, что последовательность случайных величин. Отметим, что это объект для нас относительно знакомый, поэтому мы будем часто привлекать его в качестве примера, в рамках которого вводимые определения, возможно, будет проще понять. Аналогия с последовательностью случайных величин позволяет здесь рассмотреть еще одну широко распространенную интерпретацию случайного процесса. Как уже было упомянуто в предыдущем разделе, его можно представлять себе как семейство случайных величин, проиндексированное подмножеством (или, возможно, всей) числовой прямой, – при каждом $t \in R$ имеем новую случайную величину.

Если мы фиксируем $\omega \in \Omega$, то полученная неслучайная функция $\xi(\omega, \cdot)$ называется *реализацией* случайного процесса. Наряду с этим термином употребляются также названия *траектория*, *выборочная функция*. Эту функцию далее будем обозначать просто $\xi(t)$, что не должно привести ни к каким противоречиям, хотя предложенное сейчас обозначение совпадает с обозначением случайного процесса в целом. Постараемся при пользовании этим обозначением предварительно давать пояснение – который из смыслов оно имеет.

Рассмотрим для примера снятие электрокардиограммы у некоторого пациента. В этом случае $\xi(t)$ как случайный процесс представляет собой семейство всех возможных кардиограмм, каждой из которой приписана некоторая вероятность ее появления, а $\xi(t)$ как

реализация – та конкретная кривая, которая выйдет в итоге из-под пера самописца.

Вместо общепринятых в теории случайных величин простейших числовых характеристик в теории случайных процессов используются две следующие неслучайные функции. Функция

$$m(t) = \mathbf{M}\xi(t)$$

называется *математическим ожиданием* случайного процесса $\xi(t)$. В свою очередь, функция двух аргументов

$$K(t, s) = \mathbf{cov}(\xi(t), \xi(s)) = \mathbf{M}\xi(t)\xi(s) - m(t)m(s)$$

называется *ковариационной функцией* случайного процесса. Иногда в литературе можно встретить для этой же характеристики не вполне корректный термин *корреляционная функция*. Хотя, возможно, имеет некоторый смысл рассмотрение под подобным названием функции, равной коэффициенту корреляции между значениями случайного процесса в двух точках:

$$\rho(t, s) = \rho(\xi(t), \xi(s)),$$

но мы далее условимся два названия считать синонимами.

Лемма 1. *Ковариационная функция случайного процесса обладает следующим свойством неотрицательной определенности: при произвольном натуральном k*

$$(\forall c_1, \dots, c_k \in \mathbb{R}) (\forall t_1, \dots, t_k) \sum_{i=1}^k \sum_{j=1}^k c_i c_j K(t_i, t_j) \geq 0. \quad (2.1)$$

Доказательство. Заметим, что, при произвольном выборе констант c_1, \dots, c_k и моментов времени $t_1, \dots, t_k \in T$,

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^k c_i c_j K(t_i, t_j) &= \sum_{i=1}^k \sum_{j=1}^k c_i c_j \mathbf{cov}(\xi(t_i), \xi(t_j)) = \\ &= \mathbf{M} \left(\sum_{i,j=1}^k c_i c_j (\xi(t_i) - m(t_i)) (\xi(t_j) - m(t_j)) \right) = \\ &= \mathbf{M} \left(\sum_{j=1}^k c_j (\xi(t_j) - m(t_j)) \right)^2 \geq 0. \end{aligned}$$

Название свойства неотрицательной определенности перекочевало в нашу теорию из теории матриц. Напомним, что квадратная матрица A порядка k называется там *неотрицательно определенной*, если для произвольного вектора $\vec{t} = (t_1, \dots, t_k)$ следующее скалярное произведение является неотрицательным:

$$\langle A\vec{t}, \vec{t} \rangle = \sum_{i,j=1}^k A_{i,j}t_it_j \geq 0.$$

Сравнивая это неравенство с (2.1), видим, что аналогия тут полная, достаточно положить

$$A_{i,j} = K(t_i, t_j), \quad i, j = 1, \dots, k.$$

Ниже мы увидим, что это соотношение нам поможет для простейшего способа задания гауссовского процесса.

Теперь обратимся к важнейшему в теории случайных процессов понятию его конечномерных распределений. Условимся через \bar{t} обозначать произвольное конечное подмножество T . Если при этом мы захотим подчеркнуть, что это множество состоит из n элементов, то будем писать $\bar{t}(n)$. Если подмножество $\bar{t} = \bar{t}(n)$ зафиксировано, то через $R^{\bar{t}}$ будем обозначать, как это уже объяснялось чуть выше, n -мерное евклидово пространство, отождествляемое с множеством всех действительных функций, заданных на \bar{t} . Систему этих конечномерных пространств при всевозможных \bar{t} будем считать организованной так, как это было описано в начале настоящего раздела. Тогда, в частности, $R^{\bar{s}}$ оказывается вложенным в $R^{\bar{t}}$ при $\bar{s} \subset \bar{t}$. Будем также считать все такие пространства вложенными аналогичным образом в $R^T = \{f \mid f : T \rightarrow R\}$.

Введем обозначение $\xi_{\bar{t}} = (\xi(t_1), \dots, \xi(t_n))$ для заданного $\bar{t} = \{t_1, \dots, t_n\}$ и произвольного отображения $\xi : T \rightarrow R$.

Пусть $\xi(t)$ – случайный процесс. Распределения векторов $\xi_{\bar{t}}$, когда \bar{t} пробегает все конечные подмножества множества T , называют *конечномерными распределениями* случайного процесса ξ . Если при рассмотрении конечномерных распределений мы ограничимся только случаем подмножеств T , состоящих из n элементов, то полученные распределения принято называть *n -мерными распределениями случайного процесса*.

Изучение конечномерных распределений важно, поскольку при решении практических задач мы можем одновременно работать

только со значениями случайного процесса в конечном числе точек. В частности, так всегда бывает при решении задач в рамках статистики случайных процессов. Основная проблема здесь, конечно же, в формальном обосновании возможности перенесения выводов, полученных при изучении конечномерных распределений, на случайный процесс в целом.

Как известно из курса теории вероятностей, конечномерные распределения задаются формулами

$$\left(\forall B \in \mathcal{B}(R^{\bar{t}}) \right) \mathbf{P}_{\bar{t}}(B) = \mathbf{P}(\xi_{\bar{t}} \in B).$$

Пусть $\bar{t} \supset \bar{s}$, $\pi_{\bar{t}, \bar{s}}$ – естественная проекция $R^{\bar{t}}$ на $R^{\bar{s}}$. Тогда во введенных обозначениях справедливо

$$(\forall A \in \mathcal{B}(R^{\bar{s}})) \mathbf{P}_{\bar{s}}(A) = \mathbf{P}_{\bar{t}}(\pi_{\bar{t}, \bar{s}}^{-1}(A)). \quad (2.2)$$

Это условие называется *условием согласованности* семейства распределений $\{\mathbf{P}_{\bar{t}}, \bar{t} \subset T\}$. Оказывается, оно является необходимым и достаточным для того, чтобы рассматриваемое семейство являлось семейством конечномерных распределений некоторого случайного процесса, заданного на T . Об этом – следующая основополагающая для теории случайных процессов теорема, доказанная А.Н.Колмогоровым.

Теорема 1. (Колмогоров) Пусть семейство вероятностных распределений удовлетворяет условию согласованности (2.2). Тогда найдется некоторое вероятностное пространство $\langle \Omega, \mathcal{F}, \mathbf{P} \rangle$ и случайный процесс $\xi(t)$ на нем, такой, что $\mathbf{P}_{\bar{t}}$ есть распределение $\xi_{\bar{t}}$ для произвольного конечного $\bar{t} \subset T$.

Изложим здесь идею доказательства. Полное доказательство можно найти в [4, приложение 2].

Пусть $\Omega = R^T$, $\xi(t, \omega) = \omega(t)$. Множество $B \subset R^T$ назовем цилиндрическим, если оно имеет вид $\{\omega | (\omega(t_1), \dots, \omega(t_n)) \in A\}$ для некоторого набора t_1, \dots, t_n точек T и какого-нибудь $A \in \mathcal{B}(R^n)$. Введем \mathcal{C} – класс всех цилиндрических подмножеств R^T , $\mathcal{F} = \sigma(\mathcal{C}) = \mathcal{B}(R^T)$. Для $B \in \mathcal{C}$ зададим

$$\mathbf{P}(B) = \mathbf{P}_{\bar{t}}(A).$$

Для завершения доказательства осталось проверить условия теоремы Каратеодори о продолжении вероятностной меры. Тогда, продолжая построенную вероятность \mathbf{P} на всю σ -алгебру \mathcal{F} , мы получим, что, будучи перенесенным на вероятностное пространство

$\langle \Omega, \mathcal{F}, \mathbf{P} \rangle$, процесс $\xi(t)$ обладает нужным нам семейством конечномерных распределений.

Действительно, множество $\{\xi_{\bar{t}} \in A\}$ по определению ξ является цилиндрическим, а, следовательно,

$$\mathbf{P}(\xi_{\bar{t}} \in A) = \mathbf{P}_{\bar{t}}(A)$$

для произвольного $A \in \mathcal{F}$.

Иногда полезным является вариант условия согласованности (2.2) в терминах характеристических функций. Пусть, как и раньше, \bar{t} – конечное подмножество T . Определим

$$\langle \lambda, x \rangle_{\bar{t}} = \sum_{\{k: t_k \in \bar{t}\}} \lambda(t_k) x(t_k), \quad \lambda, x \in R^T.$$

Тогда характеристическая функция распределения $\mathbf{P}_{\bar{t}}$ – это

$$\varphi_{\bar{t}}(\lambda) = \int_{R^T} \exp\{i \langle \lambda, x \rangle_{\bar{t}}\} d\mathbf{P}(x).$$

Условие согласованности распределений во введенных обозначениях эквивалентно при этом

$$(\forall \lambda \in R^T) (\bar{s} \subset \bar{t}) \Rightarrow (\varphi_{\bar{s}}(\lambda) = \varphi_{\bar{t}}(\pi_{\bar{t}, \bar{s}}(\lambda))). \quad (2.3)$$

Здесь $(\pi_{\bar{t}, \bar{s}}(\lambda))(u) = \lambda(u)$ при $u \in \bar{s}$ и равно нулю при $u \in \bar{t} \setminus \bar{s}$.

Глава 3

Цели курса и элементы его траектории

Определение случайного процесса дано. Построены и описаны также те математические конструкции, в рамках которых должна далее развиваться теория. Как же дальше предполагается продвигаться в изучении случайных процессов?

Прежде всего понятно, что нужно подробно изучить способы их корректного задания. Это связано с тем, что даже в курсах теории вероятностей при столкновении с не очень просто устроенными случайными величинами для их корректного задания приходится вводить достаточно сложное понятие функции распределения. В отличие от интуитивно прозрачного понятия ряда распределения тут приходится давать специальные пояснения, чтобы разобраться с содержанием этого нового понятия. Случайный процесс же представляет собой целое, вообще говоря, несчетное семейство случайных величин, а, следовательно, трудности процедуры его корректного задания вырастают экспоненциально.

Итак, первая проблема, которую мы будем рассматривать, будет состоять в изучении способов задания случайного процесса. Один из таких способов был, в сущности, уже намечен через теорему Колмогорова. Это задание процесса через семейство его конечномерных распределений. Разумеется, тут остаются свои сложности. Но окажется, что таким образом довольно просто задавать, например, так называемые гауссовские случайные процессы.

А что, если ограничиться только заданием математического ожидания и ковариационной функции? Конечно же, описание случайного процесса только с помощью этих двух характеристик оставляет слишком широкие возможности варьировать остальные характеристики и свойства этого процесса. Но оказалось, что для решения довольно большого разнообразия практически важных задач такой способ допустим. Те определения и свойства случайных процессов, которые налагают какие-либо требования только на их математические ожидания и ковариационные функции, составляют часть теории случайных процессов, называемую *теорией случайных процессов в широком смысле*.

При рассмотрении вопросов этой части теории весьма полезным оказывается представление случайного процесса каноническим разложением. Дадим определение. Пусть $\varphi_n(t)$, $n \in \mathbf{N} \cup \{0\}$ – система неслучайных функций, X_n , $n \in \mathbf{N}$ – некоррелированные случайные величины с нулевыми математическими ожиданиями. Тогда запись случайного процесса $\xi(t)$ в виде

$$\xi(t) = \varphi_0(t) + \sum_{n=1}^{\infty} X_n \varphi_n(t)$$

называют его *каноническим разложением*. Из самого определения немедленно получаем

$$\mathbf{M}\xi(t) = \varphi_0(t); \quad K(t, s) = \sum_{n=1}^{\infty} \sigma_n^2 \varphi_n(t) \varphi_n(s),$$

где $\sigma_n^2 = \mathbf{D}X_n$, $n \in \mathbf{N}$.

Как видим, знание канонического разложения упрощает вычисления основных характеристик случайного процесса. Следует, видимо, все же сказать о том, что каноническое разложение имеет тот же смысл, что разложение, скажем, некоторого вектора по базису векторного пространства. В этом смысле система неслучайных функций $\varphi_n(t)$, $n \in \mathbf{N} \cup \{0\}$ представляет собой координаты процесса в базисе X_n , $n \in \mathbf{N} \cup \{0\}$, где X_0 – тождественная единица.

В рамках изучаемой части теории можно дать определения практически любого функционального свойства в широком смысле. Например, случайный процесс интегрируем в широком смысле, если его математическое ожидание и ковариационная функция интегрируемы, непрерывен – если непрерывны и т.п. При этом эти свойства

будут означать интегрируемость, непрерывность и тому подобное в смысле даваемых ниже в соответствующих разделах определений лишь для процессов, имеющих нормальные (гауссовские) конечномерные распределения. Об этих процессах речь пойдет в следующем разделе. Для иных процессов требования выполнения функциональных свойств «в узком смысле» будут более жесткими. Тем самым, определения «в широком смысле» содержат меньше строгих требований, допускают несколько большую свободу, и именно поэтому так и называются.

Следующая задача курса связана с тем, что произвольный случайный процесс является, хотя и специфической, но все же функцией. Поскольку теория функций в рамках классической математики получила весьма существенное развитие, то естественно попытаться перенести хотя бы часть функциональных понятий математического анализа на случайные процессы. К этим понятиям, в частности, относятся пределы, производные, интегралы и т.п.

И, наконец, задача, которая является специфической именно для нашей теории. Видимо, чаще всего возникающей проблемой при работе со случайными процессами, является задача их прогноза, попытки угадать их дальнейшее развитие после наблюдений за их поведением в течение некоторого времени. Мы уделим решению этой задачи довольно много внимания. К сожалению, единых рецептов здесь дать не получится, и, даже в довольно простых ситуациях, эта задача окажется крайне сложной.

Некоторое время и усилия мы также потратим на изучение разных классов случайных процессов, которые возникают из потребностей практики, а также математическим моделям с привлечением этих конкретных видов процессов.

Глава 4

Гауссовские случайные процессы

Вынесенный в заголовок класс случайных процессов принято сравнивать по их месту в нашей теории с нормальными распределениями в курсах теории вероятностей и математической статистики. И, хотя связь между этими двумя понятиями, безусловно, есть, причем самая прямая, но сравнивать роли тут вряд ли корректно. Слишком разнообразен класс гауссовских случайных процессов, но и слишком часто встречаются в тех же инженерных задачах процессы, которые невозможно или крайне трудно свести к гауссовским. Но вот задать гауссовский процесс оказывается настолько же легко, как нормальное распределение. Для этого так же достаточно задания характеристик первого и второго порядка. Перейдем к определениям.

Случайный процесс $\xi(t)$ называется *гауссовским*, если все его конечномерные распределения нормальны. Таким образом, характеристическая функция любого конечномерного распределения гауссовского процесса имеет вид

$$\varphi(\vec{\lambda}) = \exp \left\{ i \langle \vec{\lambda}, \vec{a} \rangle - \frac{1}{2} \langle B\vec{\lambda}, \vec{\lambda} \rangle \right\},$$

где \vec{a} – вектор математических ожиданий, а B – ковариационная матрица координат соответствующего вектора. Как известно из курса теории вероятностей (и следует из приведенной выше формулы),

для задания нормального распределения в конечномерном случае достаточно задать \vec{a} и матрицу ковариаций B .

Для обоснования того факта, что, задав произвольный вектор математических ожиданий и любую симметричную неотрицательно определенную матрицу размерности n , мы действительно зададим предыдущим соотношением некоторое вероятностное распределение на n -мерном евклидовом пространстве, в курсе теории вероятностей просто убеждаются, что определенный n -мерный интеграл от соответствующего обратного преобразования Фурье, которое имеет вид

$$p(\vec{x}) = \sqrt{\frac{1}{(2\pi)^n |B|}} \exp \left\{ -\frac{1}{2} \langle B^{-1}(\vec{x} - \vec{a}), (\vec{x} - \vec{a}) \rangle \right\}$$

по всему R^n равен 1. Отсюда следует, что это действительно плотность некоторого распределения. Если же B имеет нулевой определитель, то находят базис из ее собственных векторов. Взяв те из них, которым соответствуют ненулевые собственные числа, с помощью аналогичной формулы обращения задают плотность распределения лишь на подпространстве, натянутом эти векторы, вне него принимая плотность равной нулю. Такое нормальное распределение называют вырожденным, и оно тоже допустимо.

Для гауссовских случайных процессов справедлив практически аналогичный результат.

Теорема 2. *Для произвольной функции $a(t)$ и функции двух переменных $K(t, s)$, удовлетворяющей условию (2.1) неотрицательной определенности, существует гауссовский случайный процесс, математическое ожидание которого $a(t)$, а ковариационная функция $K(t, s)$.*

Доказательство. Заметим, что достаточно построить процесс с тождественно нулевым математическим ожиданием, а затем прибавить к нему неслучайную функцию $a(t)$. Зафиксируем $\vec{t} = \{t_1, \dots, t_n\}$, и пусть $\mathbf{P}_{\vec{t}}$ – нормальное распределение в R^n с нулевым средним и ковариационной матрицей $B = B_{\vec{t}}$, элементы которой вычислены по правилу $B_{i,j} = K(t_i, t_j)$. Получившаяся матрица будет неотрицательно определена в силу условия (2.1). Таким образом, описанное построение возможно. Чтобы завершить доказательство, достаточ-

но заметить, что

$$\varphi_{\bar{i}}(\vec{\lambda}) = \exp \left\{ -\frac{1}{2} \langle B\vec{\lambda}, \vec{\lambda} \rangle \right\} -$$

характеристическая функция построенного распределения и проверить условия согласованности получившихся распределений (например, в форме (2.3)).

Действительно,

$$\begin{aligned} \varphi_{\bar{i}}(\pi_{\bar{i}\bar{s}}(\lambda)) &= \exp \left\{ -\frac{1}{2} \sum_{i,j \in \bar{i}} B_{i,j} \pi_{\bar{i}\bar{s}}(\lambda)(t_i) \pi_{\bar{i}\bar{s}}(\lambda)(t_j) \right\} = \\ &= \exp \left\{ -\frac{1}{2} \sum_{i,j \in \bar{s}} B_{i,j} \lambda(t_i) \lambda(t_j) \right\} = \varphi_{\bar{s}}(\lambda), \end{aligned}$$

поскольку все «лишние» слагаемые, в которых хотя бы одно из t_i, t_j не попадает в \bar{s} , на последнем шаге обращаются в 0, и теорема доказана.

Простым следствием доказанной теоремы является то, что условие неотрицательной определенности функции двух аргументов K является необходимым и достаточным для того, чтобы она являлась бы ковариационной функцией какого-либо случайного процесса.

Очень важным примером гауссовского процесса для нас является *винеровский процесс*, который служит моделью броуновского движения.

Винеровским процессом $w(t)$ мы будем называть случайный процесс, определенный при $t \geq 0$ и такой, что $w(0) = 0$ (процесс «выходит из нуля»), $w(t) - w(s)$ имеет нормальное распределение $N(0, t - s)$, если $t > s$, и обладает тем свойством, что при произвольном выборе $0 \leq t_1 < t_2 < t_3 < t_4$ случайные величины $w(t_2) - w(t_1)$ и $w(t_4) - w(t_3)$ независимы. Позднее мы скажем более точно, что этот процесс обладает независимыми приращениями. Если и дальше забегать вперед, то $w(t)$ является также и стационарным, что показывает второе из выдвинутых в определении условий.

Найдем конечномерные распределения винеровского процесса. Для этого зафиксируем $t_1, \dots, t_n \geq 0$ так, что $t_1 < \dots < t_n$ и рассмотрим вектор $\vec{w} = (w_1, \dots, w_n)$, где $w_j = w(t_j)$, $j = 1, 2, \dots, n$. Заметим, что плотность распределения вектора $X = (w_1, w_2 - w_1, \dots, w_n - w_{n-1})$ легко вычислить, т.к. его координаты – независимые нор-

мально распределенные случайные величины:

$$p_X(x_1, \dots, x_n) = \prod_{i=1}^n \left([2\pi(t_i - t_{i-1})]^{-1/2} \exp \left\{ -\frac{x_i^2}{2(t_i - t_{i-1})} \right\} \right),$$

где $t_0 = 0$. При этом $\vec{w} = AX$, где матрица A имеет вид

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}.$$

Как известно, в результате невырожденного линейного преобразования случайного вектора плотность его распределения изменяется по следующему закону

$$p_{\vec{w}}(X) = |A|^{-1} p_X(A^{-1}X).$$

Легко проверить, что $|A| = 1$, и обратная матрица двухдиагональна

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix},$$

откуда $A^{-1}X = (x_1, x_2 - x_1, \dots, x_n - x_{n-1})$, и

$$p_{\vec{w}}(x_1, \dots, x_n) = \prod_{i=1}^n \frac{\exp\{-(x_i - x_{i-1})^2/2(t_i - t_{i-1})\}}{\sqrt{2\pi(t_i - t_{i-1})}},$$

где для упрощения формулы выбрано $x_0 = 0$. Очевидно, что после раскрытия всех скобок в показателях экспонент для некоторой константы C , зависящей от t_1, \dots, t_n , мы получим

$$p_{\vec{w}}(x_1, \dots, x_n) = C \exp\{-Q(x_1, \dots, x_n)/2\},$$

причем Q – неотрицательно определенная квадратичная форма своих аргументов. Если W – матрица квадратичной формы Q , то полученное соотношение представляет собой плотность нормального распределения со средним $\vec{0}$ на пространстве размерности $\text{rang}(W)$.

Итак, конечномерные распределения винеровского процесса являются нормальными, и рассматриваемый процесс действительно гауссовский.

Кроме винеровского процесса $w(t)$ (называемого иногда броуновским движением, что, на самом деле, не вполне точно), в приложениях иногда встречается так называемый *броуновский мост* – процесс, задаваемый равенством

$$w^0(t) = w(t) - tw(1), \quad t \in [0, 1].$$

Очевидно, что $w^0(0) = w^0(1) = 0$ с вероятностью 1, а значит, начинаясь в 0, наш процесс оказывается на нулевом уровне и при $t = 1$, чем отчасти объясняется его название. Вычислим ковариационную функцию броуновского моста. Пусть $t \geq s$.

$$K(t, s) = \mathbf{M}(w(t) - tw(1))(w(s) - sw(1)) = \mathbf{M}w(t)w(s) - \\ - t\mathbf{M}w(1)w(s) - s\mathbf{M}w(1)w(t) + st\mathbf{M}w^2(1).$$

Заметим, что для произвольного a $\mathbf{M}w^2(a) = a$, а при $h > u$

$$\mathbf{M}w(h)w(u) = \mathbf{M}(w(h) - w(u))w(u) + \mathbf{M}w^2(u) = u,$$

откуда

$$K(t, s) = s - ts - st + st = s(1 - t), \quad t \geq s.$$

Выбросом случайного процесса за уровень a называют событие $\{\xi(t) > a\}$. Пусть T_a – время пребывания процесса над уровнем a за время T , N_a – количество выбросов процесса за уровень в течение этого времени, τ_a – длительность одного такого выброса. Выпишем здесь для справки средние значения введенных величин для стационарного гауссовского процесса. Понятие стационарности будет введено ниже, сейчас же договоримся под этим понимать тот факт, что распределение процесса $\xi(t)$ одно и то же при произвольном t . Подробное изложение теории выбросов, включая доказательство приводимых ниже формул, см. [5].

$$\mathbf{M}\tau_a = \pi \frac{\sigma_x}{\sigma_v} \exp \left\{ -\frac{(a - m_x)^2}{2\sigma_x^2} \right\} \left(1 - \Phi \left(\frac{a - m_x}{\sigma_x} \right) \right),$$

где $m_x = \mathbf{M}\xi(t)$, $\sigma_x^2 = \mathbf{D}\xi(t)$ (не зависят от t в силу стационарности процесса), а σ_v^2 – дисперсия производной процесса. Далее, $\mathbf{M}T_a = T\mathbf{M}\tau_a$, и наконец,

$$\mathbf{M}N_a = \frac{T\sigma_v}{2\pi\sigma_x} \exp \left\{ -\frac{(a - m_x)^2}{2\sigma_x^2} \right\}.$$

Глава 5

Некоторые важные классы случайных процессов

В этом разделе мы рассмотрим некоторые классы случайных процессов, которые традиционно выделяются в практических задачах. Эти классы пересекаются, и их пересечение, как оказывается, состоит из процессов, которые задать ничуть не сложнее, чем гауссовские. Перейдем к определениям.

Пусть $T = [a, b]$. Процесс $\xi(t)$, $t \in T$ называется *процессом с независимыми приращениями*, если для любого k и любых $t_1 < \dots < t_k \in T$ случайные величины

$$\xi(a), \xi(t_1) - \xi(a), \xi(t_2) - \xi(t_1), \dots, \xi(b) - \xi(t_k)$$

являются независимыми.

Поработаем с конечномерными распределениями. Пусть

$$a = t_0 < t_1 < \dots < t_n -$$

упорядоченная цепочка элементов индексного множества T . Обозначим через \mathbf{P}^a распределение $\xi(a)$, через $\mathbf{P}_{s,t}$ распределение приращения $\xi(t) - \xi(s)$, $t > s$. Из формулы

$$\xi(t_j) = \xi(a) + \sum_{i=1}^j (\xi(t_i) - \xi(t_{i-1}))$$

и того, что все слагаемые в выписанной сумме являются независимыми, вытекает, что, считая все распределения $\mathbf{P}_{s,t}$, $t > s$ известными, мы можем восстановить распределение каждой из случайных величин $\xi(t_j)$.

Следовательно, и распределение $\xi_{\bar{t}}$ – совместное распределение случайных величин $\xi_0 + \xi_1$, $\xi_0 + \xi_1 + \xi_2, \dots, \xi_0 + \dots + \xi_n$, где $\xi_0 = \xi(a)$, $\xi_i = \xi(t_i) - \xi(t_{i-1})$ при $i \geq 1$ можно восстановить подобно тому, как это было сделано при выводе конечномерных распределений винеровского процесса. Итак, сославшись на теорему Колмогорова, мы видим, что для задания процесса с независимыми приращениями достаточно знать распределения всех его приращений а также так называемое «стартовое» распределение \mathbf{P}^a .

Говорят, что $\xi(t)$ – *стационарный процесс*, если ни одна его вероятностная характеристика не меняется со временем. В частности, распределение случайной величины $\xi(t)$ одно и то же для всех t , для произвольных s, t, h распределения $\xi(t) - \xi(s)$ и $\xi(t+h) - \xi(s+h)$ совпадают и т.д. Если известно только, что математическое ожидание процесса константа, а также то, что для произвольных t, s, h его ковариационная функция удовлетворяет соотношению

$$K(t+h, s+h) = K(t, s),$$

то такой процесс называют стационарным в широком смысле. Понятие стационарного случайного процесса является естественным обобщением понятия последовательности одинаково распределенных случайных величин, но содержит в себе несколько больше требований. Например, здесь мы требуем, чтобы не только распределения самих случайных величин были бы одинаковы, но и то, чтобы совместные распределения ансамблей этих величин зависели бы только от количества величин ансамбля, а не от того, какие конкретно это величины.

Процесс $\xi(t)$ называется *стохастически непрерывным*, если при $s \rightarrow t$ выполняется $\xi(s) \stackrel{\mathbf{P}}{\longrightarrow} \xi(t)$.

Оказывается, если случайный процесс попадает во все три определенных выше класса, то для того, чтобы его задать, достаточно знать существенно меньше распределений, чем для общего процесса с независимыми приращениями.

Теорема 3. Пусть процесс $\xi(t)$ стохастически непрерывен, стационарен и имеет независимые приращения. Тогда для задания его

конечномерных распределений достаточно кроме \mathbf{P}^a задать только одно распределение. Этим распределением может быть любое из распределений $\mathbf{P}_{s,t}$.

Доказательство. Заметим, что можно без ограничения общности считать, что $a = 0$. Иначе просто сдвинем все аргументы процесса на одно и то же значение a , что не приведет к изменению его конечных распределений в силу заявленной стационарности. Пусть $\varphi_t(\lambda)$ – характеристическая функция распределения $\mathbf{P}_{0,t}$. В силу представления

$$\xi(t) - \xi(0) = \sum_{i=1}^n \left(\xi\left(\frac{it}{n}\right) - \xi\left(\frac{(i-1)t}{n}\right) \right)$$

и того, что распределения всех слагаемых в этой сумме одинаковы, следует соотношение

$$(\forall t, n, \lambda) \varphi_t(\lambda) = \varphi_{\frac{t}{n}}^n(\lambda). \quad (5.1)$$

Покажем далее, что ни при одном λ $\varphi_t(\lambda)$ не обращается в 0. Действительно, при $n \rightarrow \infty$ справедливо

$$\xi\left(\frac{t}{n}\right) - \xi(0) \xrightarrow{\mathbf{P}} 0,$$

откуда, по основной теореме о характеристических функциях, $\varphi_{t/n}(\lambda)$ при $n \rightarrow \infty$ стремится к тождественной единице, характеристической функции нулевой константы. Теперь требуемое следует из (5.1). Действительно, ведь если $\varphi_t(\lambda) = 0$, то для этого λ значение $\varphi_{\frac{t}{n}}(\lambda) = 0$ и к 1 стремиться не может.

Таким образом, может быть корректно определена функция $\Psi_t(\lambda)$ такая, что при всех t, λ

$$\varphi_t(\lambda) = \exp\{\Psi_t(\lambda)\}.$$

Следующей нашей целью будет доказать, что

$$(\forall t \in T)(\forall \lambda) \Psi_t(\lambda) = t\Psi_1(\lambda) \equiv t\Psi(\lambda). \quad (5.2)$$

Пусть сначала $t = \frac{1}{n}$. Выберем в (5.1) $t = 1$, получим

$$\varphi_1(\lambda) = \varphi_{1/n}^n(\lambda) \implies \exp\Psi(\lambda) = \exp\{n\Psi_{1/n}(\lambda)\},$$

откуда вытекает справедливость (5.2) для $t = \frac{1}{n}$. Запишем

$$\xi\left(\frac{m}{n}\right) - \xi(0) = \sum_{j=1}^m \left(\xi\left(\frac{j}{n}\right) - \xi\left(\frac{j-1}{n}\right) \right),$$

откуда вытекает $\varphi_{m/n}(\lambda) = \varphi_{1/n}^m(\lambda)$, что влечет равенство

$$\Psi_{\frac{m}{n}}(\lambda) = m\Psi_{\frac{1}{n}}(\lambda) = \frac{m}{n}\Psi(\lambda).$$

Мы получили желаемое для всех рациональных $t \in T$. Если, наконец, t – произвольный элемент индексного множества, рассмотрим последовательность x_n рациональных чисел, сходящуюся к t . В силу стохастической непрерывности

$$\xi(x_n) - \xi(0) \xrightarrow{\mathbf{P}} \xi(t) - \xi(0),$$

откуда $\varphi_{x_n}(\lambda) \rightarrow \varphi_t(\lambda)$, и $x_n\Psi(\lambda) \rightarrow \Psi_t(\lambda)$. С другой стороны, $x_n\Psi(\lambda) \rightarrow t\Psi(\lambda)$, и доказываемое соотношение (5.2) немедленно следует из единственности предела.

Итак, пусть мы знаем $\mathbf{P}_{0,1}$. Тогда мы знаем его характеристическую функцию, а значит и $\Psi(\lambda)$. По формуле (5.2) и заданному t восстанавливаем Ψ_t , затем φ_t , а значит, и произвольное $\mathbf{P}_{0,t}$. Теорема доказана.

В заключение приведем формулу для характеристической функции распределения вектора $\xi_{\vec{t}} = (\xi(t_1), \dots, \xi(t_n))$, вывод которой оставляем читателю в качестве упражнения, полезного для восстановления навыков в использовании свойств характеристических функций.

$$\varphi_{\vec{t}}(\vec{\lambda}) = \varphi_a(\Lambda_n) \cdot \exp \left\{ \sum_{j=1}^n (t_j - t_{j-1}) \Psi(\Lambda_n - \Lambda_{j-1}) \right\},$$

где φ_a – характеристическая функция \mathbf{P}^a ,

$$\Lambda_k = \sum_{j=1}^k \lambda_j, \quad \Lambda_0 = 0.$$

Глава 6

Понятие марковского процесса

6.1 Концепция марковости

Одним из часто применяемых обобщений процессов с независимыми приращениями является понятие марковского процесса. Оно использует представление о специфической форме зависимости приращений случайного процесса друг от друга.

Пусть $\xi(t)$ – случайный процесс, обозначим

$$\mathcal{F}_{\leq t} = \sigma(\{\xi(s), s \leq t\}), \quad \mathcal{F}_{=t} = \sigma(\xi(t)), \quad \mathcal{F}_{\geq t} = \sigma(\{\xi(s), s \geq t\}).$$

Неформально эти сигма-алгебры включают в себя все те события, о наступлении которых мы имеем полную информацию, наблюдая развитие процесса до момента t ; только в момент t ; начиная с момента t соответственно. Разумеется, практический смысл имеют только первая и вторая сигма-алгебры. Третья является лишь воображаемым (в принципе допустимым) объектом, и вводится только для того, чтобы дать строгое определение.

Процесс $\xi(t)$ называют *марковским*, если

$$(\forall t)(\forall B \in \mathcal{F}_{\geq t}) \mathbf{P}(B/\mathcal{F}_{\leq t}) = \mathbf{P}(B/\mathcal{F}_{=t}). \quad (6.1)$$

Фиксация некоторой сигма-алгебры означает, что нам известно все о событиях, ее составляющих. Подробнее об этом сказано в разделе 9. При этом $\mathcal{F}_{\leq t}$ олицетворяет собой прошлое и настоящее, $\mathcal{F}_{=t}$

6.1. Концепция марковости

– настоящее, а $\mathcal{F}_{\geq t}$ – будущее и настоящее. Таким образом, марковское свойство можно интерпретировать так, что будущее развитие процесса, по крайней мере, его вероятностные характеристики, при фиксированном настоящем не зависит от прошлого. Если учесть, что процесс с независимыми приращениями по определению обладает этим свойством, поскольку $\xi(s) - \xi(t)$ не зависит от $\xi(t) - \xi(u)$ при $u < t < s$ даже без фиксирования $\xi(t)$, то становится ясно, что понятие марковского процесса обобщает понятие процесса с независимыми приращениями.

Если выражаться совсем простым языком, то для прогнозирования поведения марковского процесса в будущем нам не требуется никакой информации о том, как он вел себя в прошлом, достаточно знать его значение в текущий момент.

Множество $\{\xi(t), t \in T\}$ называют *множеством состояний* марковского процесса. Если это множество не более, чем счетно, то марковский процесс называют *цепью Маркова*, а процесс изменения $\xi(t)$ с течением времени – переходом от состояния к состоянию. Для цепей Маркова состояния нумеруются и отождествляются со своими номерами, то есть считают, что цепь Маркова может пребывать в одном из состояний 1, 2, Марковское свойство (6.1) здесь может быть переписано следующим образом. Для произвольного натурального k и любого набора $n_1, \dots, n_{k+1} \in \mathbf{N}$ при $t_1 < \dots < t_{k+1}$, выполнено

$$\begin{aligned} \mathbf{P}(\xi(t_{k+1}) = n_{k+1} / \xi(t_1) = n_1, \dots, \xi(t_k) = n_k) &= \\ = \mathbf{P}(\xi(t_{k+1}) = n_{k+1} / \xi(t_k) = n_k) &= P_{n_k, n_{k+1}}(t_{k+1}, t_k). \end{aligned}$$

Если эта вероятность зависит только от $\Delta t = t_{k+1} - t_k$, то цепь Маркова называют *однородной*, а вероятность

$$P_{m,n}(\Delta t) = \mathbf{P}(\xi(t + \Delta t) = n / \xi(t) = m) \quad (6.2)$$

вероятностью перехода из состояния m в состояние n за время Δt . Если индексное множество счетно (чаще всего $T = \{0\} \cup \mathbf{N}$), то цепь Маркова называют *дискретной*.

Изложение теории дискретных однородных цепей Маркова и классификацию их состояний можно найти в [4], а общую теорию марковских процессов в [15]. Здесь же мы остановимся лишь на одном практически важном частном случае и сделаем ряд дополнительных предположений, фактически запрещающих значительное изменение номера текущего состояния за небольшие промежутки времени.

Пусть $\xi(t)$, $t \in T$ – однородная цепь Маркова с множеством состояний $Q = \{0, 1, 2, \dots\}$. Предположим, что

1. $(\forall n \in Q) P_{n,n+1}(\Delta t) = \lambda_n \Delta t + o(\Delta t)$;
2. $(\forall n \in Q \setminus \{0\}) P_{n,n-1}(\Delta t) = \mu_n \Delta t + o(\Delta t)$;
3. $(\forall n \in Q) P_{n,n}(\Delta t) = 1 - (\lambda_n + \mu_n) \Delta t + o(\Delta t)$;
4. $(\forall n, m \in Q) P_{n,m}(0) = \delta_{n,m}$, где $\delta_{n,m}$ – символ Кронекера, равный нулю при несовпадении индексов и единице при их совпадении.

Здесь λ_n, μ_n , $n = 0, 1, \dots$ – известные неотрицательные числовые последовательности. Позднее, при изучении процессов размножения и гибели, мы назовем их интенсивностями размножения и гибели соответственно.

Из сделанных предположений, в частности, следует, что

$$P_{m,n}(\Delta t) = o(\Delta t), \quad |n - m| \geq 2.$$

Таким образом, почти наверняка переход из состояния n за малое время возможен лишь в одно из состояний $n - 1, n, n + 1$. Цепь Маркова, удовлетворяющую выписанным предположениям, будем называть *цепью устойчивого типа* $E(\lambda_n; \mu_n)$. Среди таких цепей выделим *специальный тип*, добавив еще одно предположение: найдутся положительные числа λ, μ такие, что

$$(\forall n) \quad \lambda_n = n\lambda, \quad \mu_n = n\mu. \quad (6.3)$$

6.2 Общая модель эволюции цепи Маркова

Изучим подробнее эволюцию цепи Маркова устойчивого типа во времени. Обозначим через T_n время, в течение которого наш случайный процесс не изменяет некоторого своего фиксированного значения n (время пребывания процесса в состоянии n). Пусть $F_n(t) = \mathbf{P}(T_n \geq t)$. Тогда с учетом сделанных ранее предположений устойчивости получаем, что

$$F_n(t + \Delta t) = F_n(t)P_{n,n}(\Delta t) + o(\Delta t) =$$

6.2. Общая модель эволюции цепи Маркова

$$= F_n(t) - F_n(t)(\lambda_n + \mu_n)\Delta t + o(\Delta t),$$

откуда

$$\frac{F_n(t + \Delta t) - F_n(t)}{\Delta t} = (\lambda_n + \mu_n)F_n(t) + o(1).$$

Перейдем к пределу при $\Delta t \rightarrow 0$ и решим получившееся дифференциальное уравнение.

Учтя очевидное соотношение $F_n(0) = 1$, получим

$$F_n(t) = \exp\{-(\lambda_n + \mu_n)t\},$$

и нами доказана

Теорема 4. Для произвольного n время пребывания цепи Маркова устойчивого типа $E(\lambda_n; \mu_n)$ в состоянии n имеет экспоненциальное распределение с параметром $\lambda_n + \mu_n$:

$$\mathbf{P}(T_n < t) = 1 - e^{-(\lambda_n + \mu_n)t}, \quad t \geq 0.$$

Выведем также и уравнения, связывающие вероятности перехода устойчивой цепи Маркова из одного состояния в другое. По формуле полной вероятности при $k \geq 1$

$$\begin{aligned} P_{n,k}(t+h) &= P_{n,k-1}(t)P_{k-1,k}(h) + P_{n,k}(t)P_{k,k}(h) + \\ &\quad + P_{n,k+1}(h)P_{k+1,k}(h) + o(h), \quad n > 0, \\ P_{n,0}(t+h) &= P_{n,0}(t)P_{0,0}(h) + P_{n,1}(t)P_{1,0}(h) + o(h). \end{aligned}$$

Воспользовавшись набором наших условий 1 – 3, распишем вероятности $P_{j,k}(h)$, $j = k - 1, k, k + 1$. Деля получающиеся соотношения на h и устремляя его к 0, получим систему дифференциальных уравнений:

$$\begin{aligned} P'_{n,0}(t) &= -\lambda_0 P_{n,0}(t) + \mu_1 P_{n,1}(t), \\ P'_{n,k}(t) &= \mu_{k+1} P_{n,k+1}(t) - (\lambda_k + \mu_k) P_{n,k}(t) + \\ &\quad + \lambda_{k-1} P_{n,k-1}(t), \quad k > 0. \end{aligned}$$

Эта система носит название *системы прямых дифференциальных уравнений Колмогорова*. Аналогичным образом можно получить *систему обратных дифференциальных уравнений Колмогорова*:

$$P'_{0,k}(t) = \lambda_0 P_{0,k}(t),$$

$$P'_{n,k}(t) = \mu_n P_{n-1,k}(t) - (\lambda_n + \mu_n) P_{n,k}(t) + \lambda_n P_{n+1,k}(t), \quad n > 0.$$

Разница в названиях и способах получения этих систем объясняется тем, что для обратной системы у нас изучается как бы «предыстория» попадания в состояние k , а для прямой системы нас интересуют «последствия» выхода системы из состояния n .

Две этих системы считаются основными моделями цепей Маркова. В литературе имеются их варианты для марковских процессов и с непрерывным временем. Тогда они называются *уравнениями Колмогорова – Чепмена*.

Марковские процессы можно, подобно процессам с независимыми приращениями, задавать с помощью системы вероятностей перехода и начального распределения – вероятностей процессу начать развитие при $t = 0$ в каждом из возможных состояний. Если же мы имеем дело с конечной и однородной цепью Маркова, то можно обойтись более простыми средствами. Например, задать стартовое распределение и матрицу перехода за один шаг. Теория конечных цепей Маркова прекрасно изложена в [4], и здесь ее воспроизводить заново нет никакого смысла.

6.3 Эргодические цепи Маркова

Предположим, что мы наблюдаем эволюцию цепи Маркова достаточно долгое время. Если вероятности обнаружения процесса в каждом из состояний при этом оказываются не зависящими от того, как он развивался на старте, то такую цепь принято называть эргодической, а набор соответствующих вероятностей ее стационарным распределением. Сейчас – строгое определение.

Если независимо от выбора i для каждого j определены числа

$$p_j = \lim_{t \rightarrow \infty} P_{i,j}(t),$$

причем $\sum_j p_j = 1$, то говорят, что числа p_0, p_1, \dots образуют *стационарное распределение* цепи Маркова.

Если стационарное распределение существует, то цепь Маркова называют *эргодической*. Отметим, что в основной массе практически значимых случаев оно существует.

6.3. Эргодические цепи Маркова

Теорема 5. *Если в конечной цепи Маркова некоторая степень матрицы перехода не имеет нулевых элементов, то эта цепь эргодична.*

Доказательство. Пусть для некоторого n_0

$$a = \min_{i,j} p_{i,j}(n_0) > 0$$

Определим

$$m_j^{(n)} = \min_i p_{i,j}(n), \quad M_j^{(n)} = \max_i p_{i,j}(n)$$

и заметим, что

$$m_j^{(n+1)} = \min_i \sum_t p_{i,t} p_{t,j}(n) \geq \min_i \sum_t p_{i,t} m_j^{(n)} = m_j^{(n)}$$

в силу стохастичности матрицы перехода. Поэтому последовательность минимумов не убывает по n . Полностью аналогично проверяется, что последовательность максимумов не возрастает, т.е. получено

$$(\forall n) m_j^{(n+1)} \geq m_j^{(n)}, \quad M_j^{(n+1)} \leq M_j^{(n)}.$$

Проверим, что $z_n = M_j^{(n)} - m_j^{(n)} \rightarrow 0$, что и будет означать существование нужного нам предела p_j . При этом мы только что проверили, что эта последовательность монотонна. Зафиксируем натуральное число s .

$$\begin{aligned} p_{i,j}(n_0 + s) &= \sum_t p_{i,t}(n_0) p_{t,j}(s) = \sum_t (p_{i,t}(n_0) - a p_{j,t}) p_{t,j}(s) + \\ &+ a \sum_t p_{j,t}(n_0) p_{t,j}(s) = \sum_t (p_{i,t}(n_0) - a p_{j,t}) p_{t,j}(s) + a p_{j,j}(2s). \end{aligned}$$

Но, в силу определения a ,

$$a p_{j,t} \leq p_{i,t}(n_0) p_{j,t} \leq p_{i,t}(n_0),$$

поэтому все слагаемые в сумме, оставшейся в правой части, неотрицательны, откуда

$$\begin{aligned} p_{i,j}(n_0 + s) &\geq m_j^{(s)} \sum_t (p_{i,t}(n_0) - a p_{j,t}) + a p_{j,j}(2s) = \\ &= m_j^{(s)}(1 - a) + a p_{j,j}(2s). \end{aligned}$$

Следовательно,

$$m_j^{(n_0+s)} \geq m_j^{(s)}(1-a) + ap_{j,j}(2s).$$

Аналогично выводится неравенство

$$M_j^{(n_0+s)} \leq M_j^{(s)}(1-a) + ap_{j,j}(2s).$$

Из этих двух неравенств легко следует

$$M_j^{(n_0+s)} - m_j^{(n_0+s)} \leq (M_j^{(s)} - m_j^{(s)}) \cdot (1-a).$$

Выбирая теперь в качестве s число $s+n_0$, затем $s+2n_0, \dots$ (допустим n раз), получим, что

$$M_j^{(nn_0+s)} - m_j^{(nn_0+s)} \leq (M_j^{(s)} - m_j^{(s)}) \cdot (1-a)^n,$$

т.е. стремится к 0 при $n \rightarrow \infty$. Следовательно, для подпоследовательности z_{nn_0+s} , $n \in \mathbf{N}$. «большой» последовательности z_n доказано ее стремление к 0. Остальное здесь следует из монотонности «большой» последовательности.

То, что сумма всех p_j будет равна 1, немедленно следует из соотношения

$$\sum_j p_{i,j}(n) = 1,$$

выполненного для произвольных i, n и того факта, что эта сумма состоит из конечного числа слагаемых, а значит, законна ее перестановка с предельным переходом при $n \rightarrow \infty$. Теорема доказана.

На самом деле, условие существования «полностью заполненной» степени матрицы перехода является для конечной цепи Маркова необходимым и достаточным условием ее эргодичности, как вытекает из следующей теоремы (ее доказательство может быть найдено, например, в [4]). К тому же, для конечных цепей дополнительно можно утверждать, что все p_j окажутся строго положительными (докажите!).

Теорема 6. *Конечная цепь Маркова является эргодической тогда и только тогда, когда она неразложима, не является периодичной и для некоторого ее состояния математическое ожидание времени первого возвращения в него конечно:*

$$(\exists j) \sum_{k=1}^{\infty} k \mathbf{P}(\xi(k) = j, \xi(i) \neq j, i = 1, \dots, k-1 | \xi(0) = j) < \infty.$$

6.3. Эргодические цепи Маркова

Сейчас мы займемся выводом формул для стационарного распределения для эргодической цепи Маркова устойчивого типа. Из системы прямых уравнений Колмогорова, переходя в них к пределу при $t \rightarrow \infty$, получим систему для определения чисел p_j . При этом, поскольку все функции $P_{i,j}(t)$ ограничены и имеют пределы, то их производные обязательно стремятся к 0. В результате предельного перехода получаем

$$\begin{aligned} -\lambda_0 p_0 + \mu_1 p_1 &= 0, \\ \lambda_{k-1} p_{k-1} - (\lambda_k + \mu_k) p_k + \mu_{k+1} p_{k+1} &= 0, \quad k = 1, 2, \dots \end{aligned} \quad (6.4)$$

Пусть

$$\pi_0 = 1, \quad \pi_j = \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j}, \quad j \geq 1.$$

Покажем, что решение системы (6.4) имеет вид

$$p_j = \frac{\pi_j}{\sum_k \pi_k}, \quad j \geq 0. \quad (6.5)$$

Действительно, из первого уравнения в (6.4) следует

$$p_1 = \frac{\lambda_0}{\mu_1} p_0 = \pi_1 p_0.$$

Воспользуемся методом математической индукции. Пусть при $j \leq k$ доказана формула

$$p_j = \pi_j p_0. \quad (6.6)$$

Тогда, подставляя (6.6) для $j = k, k-1$ в первое уравнение (6.4), получаем

$$p_{k+1} = \frac{\lambda_k \pi_k}{\mu_{k+1}} p_0,$$

откуда сразу же имеем $p_{k+1} = \pi_{k+1} p_0$. Итак, формула (6.6) доказана для всех $j \geq 0$. Сложим соотношения, вытекающие из (6.6) при изменении j в пределах от 0 до ∞ и учтем, что $\sum_j p_j = 1$. Получаем

$$1 = p_0 \sum_j \pi_j \Rightarrow p_0 = \frac{1}{\sum_j \pi_j}.$$

Подставляя результат в (6.6), немедленно выводим (6.5).

Глава 7

Линейные преобразования случайных процессов

Перейдем далее к перенесению основных конструкций математического анализа на класс случайных процессов. Разумеется, основными инструментами при работе с функциями служат производные и интегралы. При этом операции дифференцирования и интегрирования линейны. Но прежде нам необходимо понятие предела – без этого, корректно введенного на классе случайных процессов понятия, невозможно построение наших основных линейных операторов. Для того, чтобы вводимый предел обладал всеми привычными свойствами, придется несколько изменить класс рассматриваемых процессов и разрешить им принимать комплексные значения.

7.1 Гильбертово пространство случайных величин

Пусть $\langle \Omega, \mathcal{F}, \mathbf{P} \rangle$ – вероятностное пространство. Рассмотрим множество заданных на нем $\xi : \Omega \rightarrow \mathbf{C}$ – комплекснозначных случайных элементов. По определению, для каждого такого элемента, $\Re \xi$, $\Im m \xi$ – случайные величины. Выделим множество тех ком-

7.1. Гильбертово пространство случайных величин

плекснозначных элементов, которые имеют конечные моменты второго порядка:

$$\mathbf{L}_2(\Omega) = \mathbf{L}_2(\langle \Omega, \mathcal{F}, \mathbf{P} \rangle) = \{\xi \mid \mathbf{M}|\xi|^2 < \infty\}.$$

На $\mathbf{L}_2(\Omega)$ определим скалярное произведение и норму:

$$\langle \xi, \eta \rangle = \mathbf{M}\xi\bar{\eta}, \quad \|\xi\|_{\mathbf{L}_2(\Omega)}^2 = \langle \xi, \xi \rangle = \mathbf{M}|\xi|^2.$$

Здесь и далее черта сверху над элементом означает комплексное сопряжение. Полнота построенного пространства во введенной норме проверяется теми же методами, что доказывается полнота пространства интегрируемых с квадратом функций в функциональном анализе. Таким образом, перед нами гильбертово пространство комплекснозначных случайных элементов.

Предположим, что у нас есть комплекснозначная случайная функция, которая при фиксации t всегда превращается в элемент \mathbf{L}_2 . Эту функцию будем называть комплекснозначным случайным $\mathbf{L}_2(\Omega)$ -процессом и далее в этом разделе договоримся рассматривать только такие.

Обозначим, как и ранее, $m(t) = \mathbf{M}\xi(t)$, имея в виду, что сейчас эта функция является комплекснозначной. Далее нам часто потребуется переходить к величинам и процессам с нулевыми средними для того, чтобы формулы принимали более обозримый вид. Поэтому введем для централизованного случайного процесса обозначение

$$\xi_*(t) = \xi(t) - m(t),$$

которое далее будем использовать без каких-либо объяснений.

Очевидно, что ковариационная функция процесса должна быть в этой ситуации определена следующим образом

$$K(t, s) = \mathbf{M}\xi_*(t)\overline{\xi_*(s)},$$

а свойство неотрицательной определенности (2.1) примет вид

$$(\forall c_1, \dots, c_k \in \mathbf{C}) (\forall t_1, \dots, t_k \in T) \sum_{i=1}^k \sum_{j=1}^k c_i \bar{c}_j K(t_i, t_j) \geq 0. \quad (7.1)$$

Заметим также, что из определения следует, что

$$(\forall s, t) K(t, s) = \overline{K(s, t)}.$$

Предел случайного $\mathbf{L}_2(\Omega)$ -процесса в точке определяется стандартным для нормированных пространств способом:

$$\xi(t) \overset{[\]}{\mathbf{L}_2(\Omega)} \longrightarrow \xi \iff \|\xi(t) - \xi\|_{\mathbf{L}_2(\Omega)} \rightarrow 0$$

при $t \rightarrow a$, например.

Нам понадобится следующий довольно простой признак существования предела:

Лемма 2. Предел в $\mathbf{L}_2(\Omega)$ процесса $\xi(t)$ при $t \rightarrow t_0$ существует тогда и только тогда, когда $(\exists m) m(t) \rightarrow m$ при $t \rightarrow t_0$ в обычном смысле, и

$$(\exists K < \infty) \lim_{s, t \rightarrow t_0} K(t, s) = K.$$

Перед тем, как перейти к доказательству, подчеркнем, что условие $K < \infty$ как бы подразумевает, что K является действительным числом. На самом же деле можно доказать, что если двойной предел, фигурирующий в лемме, существует, то он обязательно имеет нулевую мнимую часть, то есть никаких дополнительных предположений, кроме явно указанных в формулировке, не делается. Попробуйте аккуратно проверить это в качестве упражнения, полезного для восстановления навыков работы с комплексными числами.

Доказательство. Необходимость. Пусть $\xi(t) \overset{[\]}{\mathbf{L}_2} \longrightarrow \xi$. Тогда, выбрав $m = \mathbf{M}\xi$, $K = \mathbf{M}\xi_*\bar{\xi}_*$, где $\xi_* = \xi - m$, получим

$$|m(t) - m| \leq \mathbf{M}|\xi(t) - \xi| \leq \sqrt{\mathbf{M}|\xi(t) - \xi|^2} \rightarrow 0.$$

А из неравенства КБШ можно извлечь, что

$$\begin{aligned} |K(t, s) - K| &\leq |K(t, s) - \mathbf{M}\xi_*(t)\bar{\xi}_*| + |\mathbf{M}\xi_*(t)\bar{\xi}_* - K| \leq \\ &\leq \sqrt{\mathbf{M}|\xi_*(t)|^2 \mathbf{M}|\xi_*(s) - \xi_*|^2} + \sqrt{\mathbf{M}|\xi_*|^2 \mathbf{M}|\xi_*(t) - \xi_*|^2} \leq \\ &\leq c(t)\|\xi_*(s) - \xi_*\|_{\mathbf{L}_2(\Omega)} + c\|\xi_*(t) - \xi_*\|_{\mathbf{L}_2(\Omega)}, \end{aligned}$$

где константа, c не зависит ни от t , ни от s . Выбрав произвольное $\varepsilon > 0$, подберем δ так, чтобы второе слагаемое было меньше $\varepsilon/2$ при $|t - t_0| < \delta$. Нетрудно проверить, что функция $\sqrt{\mathbf{M}|\xi_*(t)|^2}$ в окрестности t_0 ограничена в силу уже доказанной части утверждения. Действительно,

$$\sqrt{\mathbf{M}|\xi_*(t)|^2} = \|\xi_*(t)\|_{\mathbf{L}_2} \leq \|\xi(t) - \xi\|_{\mathbf{L}_2} + \|\xi_*\|_{\mathbf{L}_2} + |m(t) - m|.$$

7.2. Дифференцирование и интегрирование в среднем квадратическом

Таким образом, $c(t)\|\xi_*(s) - \xi_*\|_{\mathbf{L}_2}$ также может быть сделано меньше $\varepsilon/2$ при s , близких к t_0 . Мы завершили доказательство необходимости условий.

Достаточность. Справедливо следующее соотношение:

$$\|\xi(t) - \xi(s)\|_{\mathbf{L}_2(\Omega)}^2 = |m(t) - m(s)|^2 + \|\xi_*(t) - \xi_*(s)\|_{\mathbf{L}_2(\Omega)}^2. \quad (7.2)$$

Но, как вытекает из определения нормы в $\mathbf{L}_2(\Omega)$,

$$\|\xi_*(t) - \xi_*(s)\|_{\mathbf{L}_2(\Omega)}^2 = K(t, t) + K(s, s) - 2\Re K(t, s),$$

а в этом соотношении каждое из слагаемых имеет пределом число K при $t, s \rightarrow t_0$. Итак, при выполнении условий леммы, оба слагаемых в (7.2) стремятся к 0, откуда

$$\|\xi(t) - \xi(s)\|_{\mathbf{L}_2(\Omega)}^2 \rightarrow 0$$

при $(t - s) \rightarrow 0$, что означает существование нужного предела в силу полноты рассматриваемого пространства. Лемма доказана.

Отметим, что в случае сходимости $\xi(t)$ к ξ в смысле пространства \mathbf{L}_2 имеет место также сходимость по вероятности, а, следовательно, и по распределению. Это немедленно следует из неравенства П.Л.Чебышёва:

$$\mathbf{P}(|\xi(t) - \xi| \geq \varepsilon) \leq \frac{\mathbf{M}|\xi(t) - \xi|}{\varepsilon} \leq \frac{\|\xi(t) - \xi\|_{\mathbf{L}_2(\Omega)}}{\varepsilon}$$

для произвольного $\varepsilon > 0$.

7.2 Дифференцирование и интегрирование в среднем квадратическом

В дальнейшем, если $\xi(t) \overset{[}{\mathbf{L}_2(\Omega)} \rightarrow \xi$ при $t \rightarrow b$, то будем писать

$$\overset{[}{t \rightarrow b} \text{l.i.m.} \xi(t) = \xi$$

и называть этот предел *пределом в среднем квадратическом*. Положим при $h, a \in R$

$$\Delta_h \xi(a) = \frac{\xi(a+h) - \xi(a)}{h}.$$

Если

$$(\exists z \in \mathbf{L}_2(\Omega)) \quad z = \lim_{h \rightarrow 0} \Delta_h \xi(a),$$

то $\xi(t)$ называют дифференцируемым в среднем квадратическом в точке a , а z — производной процесса в среднем квадратическом в этой точке. Обозначение: $z = \frac{d\xi}{dt}(a)$. Из леммы 2 следует, что справедливы следующие необходимые и достаточные условия дифференцируемости в среднем квадратическом:

Лемма 3. *Процесс $\xi(t)$ будет дифференцируемым в среднем квадратическом в точке a тогда и только тогда, когда выполнены два условия:*

- в этой точке существует производная функции $m(t)$ в обычном смысле;
- ковариационная функция процесса имеет в этой точке обобщенную вторую производную, а именно

$$D_s D_t K(t, s)|_{t=s=a} = \lim_{h, q \rightarrow 0} \frac{K(a+h, a+q) - K(a, a+q) - K(a+h, a) + K(a, a)}{hq} < \infty.$$

Теорема 7. *Пусть для произвольного $a \in [A, B]$ определена обобщенная вторая производная $D_s D_t K(t, s)|_{t=s=a}$, тогда случайный процесс $\xi_*(t)$ дифференцируем в среднем квадратическом на $[A, B]$, существуют обычные частные производные $\frac{\partial K}{\partial t}$, $\frac{\partial K}{\partial s}$ и $\frac{\partial^2 K}{\partial t \partial s}$, причем для характеристик производной процесса имеют место соотношения*

$$\mathbf{M} \xi_*(t) \frac{d\xi_*(s)}{dt} = \frac{\partial K}{\partial s}(t, s), \quad \mathbf{M} \frac{d\xi_*(t)}{dt} \frac{d\xi_*(s)}{dt} = \frac{\partial^2 K}{\partial t \partial s}(t, s). \quad (7.3)$$

Существование обобщенной второй производной означает, вообще говоря, нечто большее, чем существование второй смешанной частной производной. Простым достаточным условием для его выполнения служит непрерывность этой смешанной производной.

Рассуждая так же, как в курсе математического анализа доказывалась теорема о независимости вида смешанной производной от порядка дифференцирования, выводим

Следствие. *Если ковариационная функция $K(t, s)$ дважды непрерывно дифференцируема, то $\xi(t)$ имеет производную в среднем квадратическом и выполнены соотношения (7.3).*

7.2. Дифференцирование и интегрирование в среднем квадратическом

Доказательство теоремы. Прежде всего заметим, что если $\mathbf{M}\xi^2 < \infty$ и существует $\lim_{h \rightarrow a} \int \eta(h) = \eta$, то существует и конечный предел

$$\lim_{h \rightarrow a} \mathbf{M} \left(\xi \cdot \overline{\eta(h)} \right) = \mathbf{M}\xi\bar{\eta}. \quad (7.4)$$

Действительно,

$$|\mathbf{M}\xi\overline{\eta(h)} - \mathbf{M}\xi\bar{\eta}|^2 \leq \mathbf{M}\xi^2 \cdot \|\eta(h) - \eta\|_{\mathbf{L}_2}^2 \rightarrow 0.$$

Поскольку $\mathbf{M}\xi_* = 0$, то первое условие леммы 3 проверять не нужно, а значит определен предел $\Delta_h \xi_*(s)$ в среднем квадратическом при $h \rightarrow 0$. Таким образом, определен и

$$\lim_{h \rightarrow 0} \mathbf{M}\xi_*(s) \overline{\Delta_h \xi_*(t)} = \lim_{h \rightarrow 0} \frac{K(s, t+h) - K(s, t)}{h},$$

который равен $\frac{\partial K}{\partial t}$. Но, с другой стороны, по аналогии с (7.4), видим, что он же равен и $\mathbf{M}\xi_*(t) \frac{d\xi_*}{dt}(s)$. Существование другой частной производной первого порядка очевидно из соображений симметрии. Далее,

$$\begin{aligned} \mathbf{M} \frac{d\xi_*}{dt}(s) \overline{\frac{d\xi_*}{dt}(t)} &= \lim_{h \rightarrow 0} \mathbf{M} \Delta_h \xi_*(s) \overline{\frac{d\xi_*}{dt}(t)} = \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\mathbf{M}\xi_*(s+h) \overline{\frac{d\xi_*}{dt}(t)} - \mathbf{M}\xi_*(s) \overline{\frac{d\xi_*}{dt}(t)} \right) = \\ &= \lim_{h \rightarrow 0} \frac{K'_t(s+h, t) - K'_t(s, t)}{h} = \frac{\partial^2 K}{\partial s \partial t}(s, t). \end{aligned}$$

Теорема доказана.

Займемся определением определенного интеграла.

Пусть случайный процесс $\xi(t)$ задан на интервале действительной прямой $[A, B]$, $A = t_0 < t_1 < \dots < t_m = B$, $\Delta t_j = t_j - t_{j-1}$. Рассмотрим

$$S_m = \sum_{j=1}^m \xi(\theta_j) \Delta t_j, \quad \theta_j \in [t_{j-1}, t_j], \quad j = 1, \dots, m.$$

Обозначим $\Delta = \max_j \Delta t_j$. Говорят, что процесс $\xi(t)$ *интегрируем в среднем квадратическом*, если найдется элемент $I \in \mathbf{L}_2$ такой, что независимо от выбора точек t_j, θ_j , подчиняющихся сформулированным выше условиям, справедливо

$$I = \Delta \rightarrow 0 \int \text{l.i.m.} S_m.$$

Обозначение, принятое для интеграла I от процесса в среднем квадратическом совпадает с обозначением обычного определенного интеграла:

$$I = \int_A^B \xi(t) dt.$$

При этом, конечно же, не следует забывать, что I является интегрируемой с квадратом случайной величиной, а не случайным процессом. Полностью аналогично предыдущей теореме можно доказать следующее утверждение.

Теорема 8. Пусть $m(t)$ интегрируема на $[A, B]$, и определен двойной интеграл $\int_A^B \int_A^B K(t, s) dt ds$, тогда процесс $\xi(t)$ интегрируем в среднем квадратическом, причем

$$\mathbf{M} \overline{\int_A^B \xi_*(t) dt} = \int_A^B K(t, s) ds,$$

$$\mathbf{M} \int_A^B \xi_*(t) dt \overline{\int_A^B \xi_*(t) dt} = \int_A^B \int_A^B K(t, s) dt ds.$$

Как обычно, заменяя верхний предел в определенном интеграле некоторым текущим $s \in [A, B]$, можно превратить результат интегрирования в случайный процесс. Полезным и весьма несложным упражнением для читателя будет приспособление приведенной только что теоремы под вычисление характеристик полученного процесса.

7.3 Стохастический интеграл неслучайной функции

Пусть U – множество, \mathcal{A} – некоторый подкласс $\mathcal{P}(U)$ (системы подмножеств множества U), \hat{m} – мера на $\sigma(\mathcal{A})$. Отображение $\mu : \sigma(\mathcal{A}) \rightarrow \mathbf{L}_2$ называется элементарной ортогональной стохастической мерой со структурной функцией \hat{m} , если:

1. $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$ при $A_1 \cap A_2 = \emptyset$;

7.3. Стохастический интеграл неслучайной функции

2. $(\forall A) \mathbf{M}\mu(A) = 0, \mathbf{M}|\mu(A)|^2 = \hat{m}(A);$
3. $\mathbf{M}\mu(A)\overline{\mu(B)} = 0$ при $A \cap B = \emptyset$ (условие ортогональности).

Рассмотрим пространство

$$\mathbf{L}_2(\hat{m}) = \left\{ \psi : U \rightarrow \mathbf{C} : \int_U |\psi(t)|^2 d\hat{m} < \infty \right\}.$$

Пусть $A_1, \dots, A_k \in \mathcal{A}$ образуют конечное разбиение U . Тогда для произвольного набора констант c_1, \dots, c_k

$$\varphi(t) = \sum_{j=1}^k c_j \mathbf{1}_{A_j}(t)$$

назовем простой функцией (здесь $\mathbf{1}_A$ – индикатор множества A , т.е. функция, равная 1 на A и 0 за его пределами). Для простой функции выписанного вида по определению положим

$$\int_U \varphi(t) d\mu = \sum_{j=1}^k c_j \mu(A_j).$$

Для произвольной неотрицательной функции $\varphi \in \mathbf{L}_2(\hat{m})$ можно, как обычно, построить последовательность простых функций, сходящихся к φ в поточечном смысле. Теми же способами, что и при построении определенного интеграла в математическом анализе, доказывается, что определение

$$\int_U \varphi(t) d\mu = \lim_{n \rightarrow \infty} \int_U \varphi_n(t) d\mu$$

здесь не зависит от выбора конкретной последовательности простых функций, лишь бы она сходилась к φ поточечно.

В общем случае определение завершается разбиением неслучайной функции из $\mathbf{L}_2(\hat{m})$ на действительную и мнимую, а затем каждой из них на положительную и отрицательную части.

Пусть $\varphi(t), \psi(t)$ – простые функции. Без ограничения общности можно считать, что

$$\varphi(t) = \sum_j f_j \mathbf{1}_{A_j}(t), \quad \psi(t) = \sum_j g_j \mathbf{1}_{A_j}(t)$$

для одного и того же набора A_j , $j = 1, \dots, n$. Тогда

$$\begin{aligned} \left\langle \int_U \varphi(t) d\mu, \int_U \psi(t) d\mu \right\rangle_{\mathbf{L}_2} &= \mathbf{M} \left(\int_U \varphi(t) d\mu \cdot \int_U \overline{\psi(t)} d\mu \right) = \\ &= \sum_{j,k} f_j \bar{g}_k \mathbf{M} \left(\mu(A_j) \overline{\mu(A_k)} \right) = \sum_k f_k \bar{g}_k \hat{m}(A_k) = \langle \varphi, \psi \rangle_{\mathbf{L}_2(\hat{m})}. \end{aligned}$$

Здесь, как обычно делается для пространств интегрируемых с квадратом относительно некоторой меры функций, мы используем скалярное произведение

$$\langle \varphi, \psi \rangle_{\mathbf{L}_2(\hat{m})} = \int_U \varphi(t) \overline{\psi(t)} \hat{m}(dt).$$

Какими бы ни были $\varphi, \psi \in \mathbf{L}_2(\hat{m})$, при помощи очевидным образом организованного предельного перехода из полученного соотношения для простых функций легко получается

$$\left\langle \int_U \varphi(t) d\mu, \int_U \psi(t) d\mu \right\rangle_{\mathbf{L}_2(\Omega)} = \langle \varphi, \psi \rangle_{\mathbf{L}_2(\hat{m})}. \quad (7.5)$$

Таким образом, доказана

Теорема 9. *Стохастический интеграл осуществляет изометрическое соответствие между $\mathbf{L}_2(\hat{m})$ и $\mathbf{L}_2(\Omega)$.*

Эта теорема позволит нам вместо случайных величин, удаленных друг от друга на определенное расстояние в пространстве $\mathbf{L}_2(\Omega)$, строить неслучайные функции, так же расположенные в пространстве $\mathbf{L}_2(\hat{m})$.

Глава 8

Спектральные представления случайных процессов

Настало время обратиться к наиболее интересной специфической задаче изучаемой теории – задаче прогнозирования случайных процессов в некоторый фиксированный момент в будущем. В основе большинства методов построения таких прогнозов с минимальной среднеквадратической ошибкой лежит переход от собственно случайного процесса к специальной неслучайной функции. Затем эта функция разлагается по функциональному базису (тригонометрической системе Фурье в построенных в предыдущих разделах пространствах), а после возвращения в пространство случайных величин путем обратного преобразования, возникает искомый прогноз. Естественно, только что данное описание является очень неточным и приблизительным, но объясняет суть тех наших действий, к которым мы сейчас приступаем.

Собственно задаче прогноза будет посвящен следующий раздел, а сейчас мы рассмотрим его технические и теоретические аспекты. Для начала разберемся с понятием спектра.

8.1 Спектральная функция

Пусть $\xi(t)$ – комплекснозначный стационарный случайный процесс, заданный на отрезке $[A, B]$, имеющий нулевое математическое ожидание. Тогда его ковариационная функция может быть преобразована следующим образом

$$K(t, s) = \mathbf{M}\xi(t)\overline{\xi(s)} = \mathbf{M}\xi(t-s+A)\overline{\xi(A)} = K(t-s),$$

то есть заменена функцией одного аргумента. Условие неотрицательной определенности (7.1) переписывается теперь в виде

$$(\forall k)(\forall c_1, \dots, c_k \in \mathbf{C})(\forall t_1, \dots, t_k) \sum_{i=1}^k \sum_{j=1}^k c_i \bar{c}_j K(t_i - t_j) \geq 0. \quad (8.1)$$

Получим сейчас некоторые следствия этого варианта условия неотрицательной определенности.

Лемма 4. Пусть выполнено (8.1). Тогда

1. $K(0)$ – неотрицательное действительное число;
2. $(\forall t) K(t) = \overline{K(-t)}$;
3. $(\forall t) |K(t)| \leq K(0)$;
4. если $K(0) = 0$, то $K(t) = 0$ тождественно.

Доказательство. 1). Взяв в (8.1) $k = 1$, $c_1 = 1$, $t_1 = 0$, получим требуемое неравенство $K(0) \geq 0$.

2). В условии неотрицательной определенности (8.1) возьмем $k = 2$, $t_1 = t$, $t_2 = 0$. Видим, что для произвольных комплексных констант c_1, c_2 выполнено

$$(|c_1|^2 + |c_2|^2)K(0) + c_1 \bar{c}_2 K(t) + \bar{c}_1 c_2 K(-t) \geq 0. \quad (8.2)$$

В силу уже проверенного факта $K(0) \in \mathbf{R}$ отсюда следует

$$(\forall a \in \mathbf{C}) aK(t) + \bar{a}K(-t) \in \mathbf{R}.$$

Пусть $a = a_1 + ia_2$, $K(t) = K_1 + iK_2$, $K(-t) = K_3 + iK_4$. Тогда, подставляя эти выражения в формулу

$$\Im(aK(t) + \bar{a}K(-t)) = 0,$$

8.1. Спектральная функция

получаем, что при любом выборе действительных чисел a_1, a_2 должно выполняться

$$a_1(K_2 + K_4) + a_2(K_1 - K_3) = 0,$$

откуда $K_2 = -K_4$, $K_1 = K_3$, что и требовалось доказать.

3). Выберем в (8.1) $k = 2$, $t_1 = t$, $t_2 = 0$, а также $c_1 = |K(t)|$, $c_2 = -K(t)$. Если учесть, что $|K(-t)| = |K(t)|$, то из (8.2) видим, что

$$2|K(t)|^2 K(0) - 2|K(t)|^3 = 2|K(t)|^2 (K(0) - |K(t)|) \geq 0.$$

Заметим, что если $K(t) = 0$, то доказываемое утверждение тривиально в силу первого, уже доказанного утверждения леммы. Если же это не так, то утверждение немедленно вытекает из выписанного неравенства после деления обеих его частей на $2|K(t)|^2$.

4). Тривиально следует из третьего утверждения. Лемма доказана.

Фундаментальным результатом для этого раздела теории случайных процессов служит возможность представления ковариационной функции стационарного случайного процесса в виде специального интеграла Лебега - Стильтьеса. Тем самым, она всегда оказывается преобразованием Фурье некоторой функции распределения, что позволяет применить к изучению ковариационных функций стационарных случайных процессов теорию характеристических функций вероятностных распределений. Сформулируем точнее.

Теорема 10. (Бохнер-Хинчин) *Непрерывная функция действительного аргумента $K(t)$, принимающая комплексные значения, удовлетворяет условию неотрицательной определенности (8.1) тогда и только тогда, когда найдется функция распределения $F(\lambda)$ такая, что*

$$(\forall t) K(t) = K(0) \int_{-\infty}^{\infty} e^{it\lambda} dF(\lambda). \quad (8.3)$$

Функция, существование которой утверждается в теореме, носит название *спектральной функции*, а представление (8.3) принято называть *спектральным представлением* ковариационной функции. Доказательство теоремы Бохнера - Хинчина можно найти в

книге [6, параграф 37]. Здесь мы ограничимся лишь формулировкой. Доказательство же приведем лишь для случая дискретного времени.

Теорема 11. *Последовательность комплексных чисел $K(n)$, $n \in \mathbf{Z}$ неотрицательно определена в смысле (8.1) тогда и только тогда, когда найдется такая функция распределения $F(\lambda)$, которое сосредоточено на интервале $[-\pi, \pi]$, что для произвольного целого n справедливо*

$$K(n) = K(0) \int_{-\pi}^{\pi} e^{in\lambda} dF(\lambda),$$

причем эта функция распределения определена однозначно.

Доказательство. Можно считать, что $K(0) = 1$ (иначе перейдем к рассмотрению $\tilde{K}(n) = K(n)/K(0)$). Если спектральное представление имеется, то

$$\begin{aligned} \sum_{j,k} c_j \bar{c}_k K(n_j - n_k) &= \int_{-\pi}^{\pi} \sum_{j,k} c_j \bar{c}_k e^{i\lambda(n_j - n_k)} dF(\lambda) = \\ &= \int_{-\pi}^{\pi} \left| \sum_j c_j e^{i\lambda n_j} \right|^2 dF(\lambda) \geq 0, \end{aligned}$$

что завершает доказательство необходимости.

Обратно. Пусть справедливо (8.1). Имея в виду формулу обратного преобразования Фурье, при $\rho \in (0, 1)$, $\lambda \in [-\pi, \pi]$ рассмотрим

$$f_\rho(\lambda) = \frac{1}{2\pi} \sum_{s \in \mathbf{Z}} K(s) \rho^{|s|} e^{-i\lambda s}.$$

Так как $|K(s)e^{-i\lambda s}| \leq K(0) = 1$, то, по признаку Вейерштрасса, выписанный ряд сходится абсолютно и равномерно относительно λ . При этом из обычных формул для преобразований Фурье при целых t имеем

$$K(t) \rho^{|t|} = \int_{-\pi}^{\pi} e^{it\lambda} f_\rho(\lambda) d\lambda \Rightarrow \int_{-\pi}^{\pi} f_\rho(\lambda) d\lambda = 1.$$

8.1. Спектральная функция

Из выписанных формул следует, что для проверки того, что функция f_ρ – плотность вероятностного распределения, достаточно проверить ее неотрицательность. Запишем (8.1), выбрав $c_n = \rho^n e^{iun}$, $t_n = -n$ и распространив суммирование до ∞ .

$$0 \leq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \rho^n e^{iun} \rho^m e^{-ium} K(m-n).$$

Это можно сделать в силу сходимости выписанного ряда. Обозначим $m-n$ за z . Тогда $m = n+z \geq 0$. Отсюда при $z > 0$ n может принимать любые неотрицательные значения, а при $z \leq 0$ обязательно $n \geq -z$. Итак, $I_1 + I_2 \geq 0$, где

$$I_1 = \sum_{z \leq 0} \sum_{n=-z}^{\infty} \rho^{m+n} e^{-iuz} K(z) = \sum_{z \leq 0} \sum_{m=0}^{\infty} \rho^{|z|+2m} e^{-iuz} K(z),$$

$$I_2 = \sum_{z=1}^{\infty} \sum_{n=0}^{\infty} \rho^{z+2n} e^{-iuz} K(z).$$

Собирая две последние формулы (с очевидным переобозначением в формуле для I_1), получаем

$$0 \leq \sum_{z \in \mathbf{Z}} \sum_{n=0}^{\infty} \rho^{|z|+2n} e^{-iuz} K(z) = 2\pi f_\rho(u) \sum_{n=0}^{\infty} \rho^{2n} = \frac{2\pi f_\rho(u)}{1-\rho^2},$$

откуда и следует неотрицательность f_ρ . Введем

$$F_\rho(\lambda) = \int_{-\pi}^{\lambda} f_\rho(u) du.$$

Понятно, что эта функция распределения имеет характеристическую функцию $\rho^{|t|} K(t)$. При $\rho \rightarrow 1$ эти функции сходятся к $K(t)$. Привлекая теорему непрерывности для характеристических функций из [7, с. 343], получаем, что F_ρ слабо сходится при соответствующем изменении ρ к F – функции распределения с характеристической функцией $K(t)$, что немедленно приводит к

$$K(t) = \int_{-\pi}^{\pi} e^{it\lambda} dF(\lambda).$$

Единственность F следует из взаимной однозначности соответствия между характеристическими функциями и функциями распределения. То, что это распределение сосредоточено на $[-\pi, \pi]$ немедленно получается из того, что интеграл от его плотности по этому отрезку равен единице. Тем самым, теорема полностью доказана.

Будем до конца раздела, если не оговорено иное, считать, что \mathcal{A} – алгебра отрезков в $[-\pi, \pi]$, если рассматривается случайная последовательность и в \mathbf{R} в непрерывном случае. Оказывается, представление ковариационной функции в виде интеграла приводит к близкой формуле и для самого случайного процесса. Правда, при этом вместо интеграла Лебега-Стилтьеса приходится использовать интеграл по стохастической мере. Спектральная же мера начинает играть роль соответствующей структурной функции.

Напомним, что процесс $\xi(t)$ считается в этом разделе центрированным, т.е. $\mathbf{M}\xi(t) = 0$. Обозначим через \mathbf{H}_ξ замыкание в $\mathbf{L}_2(\Omega) = \mathbf{L}_2(\langle \Omega, \mathcal{F}, \mathbf{P} \rangle)$ множества конечных линейных комбинаций вида $\sum_j \alpha_j \xi(t_j)$, и пусть \mathbf{P}_ξ – мера, порожденная спектральной функцией, т.е. *спектральное распределение*.

Теорема 12. *Может быть построена элементарная ортогональная стохастическая мера $\mu : \mathcal{A} \rightarrow \mathbf{H}_\xi$ со структурной функцией $m(A) = K(0)\mathbf{P}_\xi(A)$ такая, что*

$$(\forall t) \xi(t) = \int e^{itu} d\mu(u), \quad (8.4)$$

причем интеграл берется по интервалу $[-\pi, \pi]$ в случае стационарной случайной последовательности и по $(-\infty, \infty)$ в непрерывном случае.

Доказательство. Проведем полное доказательство для случайной последовательности. Построим изометрическое отображение $I : \mathbf{L}_2(m) \rightarrow \mathbf{H}_\xi$. Для начала положим

$$I \left(\sum_{j=1}^k c_j e^{it_j u} \right) = \sum_{j=1}^k c_j \xi(t_j).$$

Ясно, что для функций рассматриваемого вида выполняется свойство изометричности:

$$\langle I(e^{itu}), I(e^{isu}) \rangle_{\mathbf{L}_2(\langle \Omega, \mathcal{F}, \mathbf{P} \rangle)} = \mathbf{M}\xi(t)\bar{\xi}(s) = K(t-s),$$

8.1. Спектральная функция

$$\langle e^{itu}, e^{isu} \rangle_{\mathbf{L}_2(m)} = \int e^{itu} \overline{e^{isu}} dm(u) = K(t-s).$$

Отметим теперь, что конечные линейные комбинации функций e^{itu} всюду плотны в $\mathbf{L}_2(m)$. Действительно, в этом пространстве всюду плотно множество ограниченных функций, в нем – дважды дифференцируемых, в последнем – периодических, в котором всюду плотно множество частичных сумм сходящихся рядов Фурье, т.е. именно таких линейных комбинаций.

Продолжим I на $\mathbf{L}_2(m)$ с сохранением изометричности и линейности. Отметим, что такое продолжение строится при помощи теоремы Хана - Банаха посредством предельного перехода в процессе аппроксимации элементов $\mathbf{L}_2(m)$ соответствующими линейными комбинациями, т.е. по непрерывности.

Пусть

$$\mu(A) = I(\mathbf{1}_A), \quad A \in \mathcal{A},$$

тогда, поскольку I – линейная функция, а для непересекающихся множеств A, B справедливо $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B$, то

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Очевидно также, что в силу изометричности построенного отображения

$$(\forall A \in \mathcal{A}) \mathbf{M}|\mu(A)|^2 = \langle \mu(A), \mu(A) \rangle_{\mathbf{L}_2(\Omega)} = \int \mathbf{1}_A^2 dm = m(A).$$

Из того, что при произвольном выборе констант c_j и точек $t_j \in T$ выполняется

$$\mathbf{M}I\left(\sum c_j e^{it_j u}\right) = \mathbf{M}\left(\sum c_j \xi(t_j)\right) = 0$$

и непрерывности I немедленно следует, что для любой $f \in \mathbf{L}_2(m)$ имеет место $\mathbf{M}I(f) = 0$ и, в частности,

$$\mathbf{M}\mu(A) = \mathbf{M}I(\mathbf{1}_A) = 0.$$

Наконец, для непересекающихся множеств $A, B \in \mathcal{A}$ справедливо

$$\begin{aligned} \mathbf{M}\mu(A)\overline{\mu(B)} &= \langle \mu(A), \mu(B) \rangle_{\mathbf{L}_2(\langle \Omega, \mathcal{F}, \mathbf{P} \rangle)} = \\ &= \langle \mathbf{1}_A, \mathbf{1}_B \rangle_{\mathbf{L}_2(m)} = \int \mathbf{1}_A \mathbf{1}_B dm = m(A \cap B) = 0. \end{aligned}$$

Тем самым, μ – элементарная ортогональная стохастическая мера со структурной функцией m .

Рассмотрим последовательность разбиений

$$-\pi = a_0 < a_1 < \dots < a_N = \pi$$

такую, что $\max_{1 \leq j \leq N} \{a_j - a_{j-1}\} \rightarrow 0$ при $N \rightarrow \infty$. Ясно, что если

$$S_N(t, u) = \sum_{j=1}^N e^{ita_j} \mathbf{1}_{[a_{j-1}, a_j)}(u),$$

то $S_N(t, u) \rightarrow e^{itu}$ при $N \rightarrow \infty$ равномерно по u . В силу линейности отображения I при каждом t

$$I(S_N(t, u)) = \sum_{j=1}^N e^{ita_j} \mu([a_{j-1}, a_j)) \rightarrow \int e^{itu} d\mu(u)$$

в среднем квадратическом, но в силу непрерывности

$$I(S_N(t, u)) \rightarrow I(e^{itu}) = \xi(t).$$

Теорема теперь немедленно вытекает из единственности предела.

Элементарную ортогональную стохастическую меру, построенную при доказательстве теоремы, условимся далее называть *стохастической спектральной мерой* процесса $\xi(t)$.

Исходя из (8.4) и линейности отображения I , нетрудно вывести, что для функций, имеющих вид $\varphi(u) = \sum_{j=1}^k c_j e^{it_j u}$ справедливо

$$I(\varphi) = \int \varphi(u) d\mu(u).$$

Привлекая единственность продолжения линейной изометрии I , получающейся, как следствие теоремы Хана-Банаха, видим, что выписанное представление справедливо для произвольной функции $\varphi \in \mathbf{L}_2(m)$.

Следствие. *Стохастический интеграл*

$$f \mapsto \int f(u) d\mu(u)$$

осуществляет изометрическое соответствие между пространствами $\mathbf{L}_2(m)$ и \mathbf{H}_ξ , причем $e^{iut} \mapsto \xi(t)$.

8.2. Формула Котельникова – Шеннона

Отметим, что вместо стохастической меры в последнем представлении можно использовать случайный процесс

$$Z(u) = \mu((-\infty, u)).$$

При этом соотношение между введенным ранее спектральным представлением случайного процесса (8.4) и записью

$$\xi(t) = \int e^{itu} dZ(u)$$

такое же, как соотношение между обычной записью интеграла Лебега – Стильтеса (с привлечением функции ограниченного изменения) и записью этого интеграла как интеграла по вероятностной мере на $(-\infty, \infty)$. Процесс $Z(u)$ в последней записи обладает *ортгоналными приращениями*, т.е. при произвольном выборе чисел $a < b < c < d$ справедливо

$$\mathbf{M}(Z(b) - Z(a))\overline{(Z(d) - Z(c))} = 0.$$

Пример. Пусть A, η – неотрицательные случайные величины, φ имеет равномерное распределение на $[-\pi, \pi]$ и не зависит от A, η . Рассмотрим стационарный случайный процесс

$$\xi(t) = A \cos(\eta t + \varphi) = \frac{1}{2} A e^{i\eta t} e^{i\varphi} + \frac{1}{2} A e^{-i\eta t} e^{-i\varphi}.$$

Последние два слагаемых и представляют собой спектральное представление этого процесса. Соответствующий процесс с ортогональными приращениями постояен на каждом из интервалов $(-\infty, -\eta), (-\eta, \eta), (\eta, \infty)$. В точке $-\eta$ он совершает скачок величины $\frac{1}{2} A e^{-i\varphi}$, а в точке η – величины $\frac{1}{2} A e^{i\varphi}$. Спектральная мера μ рассматриваемого процесса $\xi(t)$ сосредоточена в двух точках $-\pm\eta$.

8.2 Формула Котельникова – Шеннона

В качестве первого примера применения спектральных представлений получим формулу, находящую свои приложения при организации систем связи. Пусть $\xi(t)$, $t \in \mathbf{Z}$ – стационарная случайная последовательность. Предположим, что при некотором натуральном $a \geq 2$ носитель спектральной меры m этой последовательности сосредоточен в $[-\frac{\pi}{a}, \frac{\pi}{a}]$. Тогда на этом носителе равномерно

(и абсолютно) сходится ряд

$$\sum_{z \in \mathbf{Z}} \alpha_z e^{izua} = e^{iut},$$

где

$$\alpha_z(t) = \frac{\sin\left(\frac{\pi t}{a} - \pi z\right)}{\left(\frac{\pi t}{a} - \pi z\right)}$$

коэффициент Фурье. Из равномерной сходимости следует сходимость в \mathbf{L}_2 , а значит имеет место представление

$$\xi(t) = \int_{-\pi}^{\pi} e^{iut} d\mu(u) = \sum_{z \in \mathbf{Z}} \alpha_z(t) \int_{-\pi}^{\pi} e^{izau} d\mu(u),$$

откуда следует, что

$$\xi(t) = \sum_{z \in \mathbf{Z}} \alpha_z(t) \xi(az).$$

Эта формула и носит название формулы Котельникова - Шеннона. Она, в частности, означает, что в этой ситуации для восстановления любого значения последовательности ξ достаточно знать, какие значения она принимала в моменты времени, кратные a .

Другое (видимо, более современное и более практичное) восприятие формулы Котельникова – Шеннона связано с принципиальной возможностью дискретизации (оцифровки, квантификации) непрерывного аналогового сигнала. Проблема состоит в том, что непрерывный сигнал имеет и непрерывный спектр. Возможно ли его неискаженное восстановление из цифровой, дискретной формы? Сделанные выше предположения о носителе меры выполняются, поскольку на практике различные радиотехнические устройства (фильтры, усилители и другие) имеют ограниченную полосу пропускания, что приводит к ограничению спектра сигнала некоторой граничной частотой.

Полученная выше формула может интерпретироваться с этой точки зрения так. Непрерывный сигнал $\xi(t)$, ограниченный по спектру частотой a , полностью определяется совокупностью мгновенных значений (наблюдений) в моменты времени t_k , отстоящие друг от друга на интервал времени, не больший, чем $\Delta t = \frac{1}{2a}$. Формула при этом переписывается, например, в виде

$$\xi(t) = \sum_{k=-\infty}^{+\infty} \xi(k\Delta t) \frac{\sin(2\pi a(t - k\Delta t))}{2\pi a(t - k\Delta t)}.$$

8.2. Формула Котельникова – Шеннона

k -й член этого ряда представляет собой отклик фильтра низких частот с постоянным коэффициентом передачи в пределах полосы частот от нуля до a на очень короткий импульс с амплитудой $\xi(k\Delta t)$.

Процесс формирования последовательности наблюдений в выбранные моменты времени называют дискретизацией непрерывного сигнала. Восстановление непрерывного сигнала осуществляется путем подачи дискретного сигнала на фильтр низких частот. Отклик фильтра на k -е наблюдение определяется выражением

$$s_k = \xi(k\Delta t) \frac{\sin(2\pi a(t - k\Delta t))}{2\pi a(t - k\Delta t)}.$$

Суммируясь, эти отклики дают на выходе фильтра низких частот исходный неискаженный сигнал.

Отметим два важных обстоятельства. Во-первых, точное восстановление сигнала имеет место только при достаточно высокой частоте дискретизации. Минимальное допустимое значение $2a$ такой частоты называют *частота Найквиста*. Обычно, на практике частоту дискретизации выбирают выше предела Найквиста. Так, например, частота Найквиста для речевого сигнала, частоты которого составляют, как правило, диапазон от 0,3 до 3,4 килогерц принимается за 6,8 килогерц, для обычных систем телевидения – 12 килогерц. Выбор частоты дискретизации для систем телевидения высокой четкости подбирается из их диапазона частот.

Во-вторых, точное восстановление сигнала возможно при суммировании бесконечного числа откликов, что соответствует сигналу, неограниченному во времени. Но в действительности, сигналы являются ограниченными и по спектру и по времени. Поэтому при теоретической возможности точного восстановления сигнала на практике неизбежно приходится отсекал члены ряда Котельникова – Шеннона с достаточно большими по модулю номерами. Это приводит к ошибкам восстановления, хотя сама формула дает теоретическую возможность делать эти ошибки сколь угодно малыми.

8.3 Интегро-дифференциальные преобразования стационарных процессов

Здесь мы рассмотрим одно из применений спектральной теории стационарных случайных процессов, которое позволит нам получать характеристики интегралов и производных от процессов в среднем квадратическом. Пусть A – линейный оператор, являющийся комбинацией линейных дифференциальных операторов и интегральных операторов с ядрами, зависящими лишь от разности аргументов. Применение таких операторов к стационарным случайным процессам в результате снова дает стационарный процесс. Получим формулу для вычисления спектральной плотности результирующего процесса.

Введем в рассмотрение спектральное представление стационарного процесса $\xi(t)$. Пусть μ – стохастическая спектральная мера этого процесса, то есть

$$\xi(t) = \int e^{it\lambda} d\mu(\lambda),$$

где интеграл без пределов по заключенному выше соглашению берется либо по отрезку $[-\pi, \pi]$, либо по всей числовой прямой в зависимости от того, на каком индексном множестве задан изучаемый случайный процесс. Напомним, что структурная функция меры μ всегда является конечной мерой на соответствующем индексном множестве.

Сначала рассмотрим оператор дифференцирования. По определению производной в среднем квадратическом,

$$\frac{d\xi}{dt}(t) = \lim_{h \rightarrow 0} \frac{\xi(t+h) - \xi(t)}{h} = \lim_{h \rightarrow 0} \int \frac{e^{i(t+h)\lambda} - e^{it\lambda}}{h} d\mu(\lambda).$$

Здесь предел можно внести под знак интеграла, поскольку стоящая под ним дробь равномерно (относительно λ) сходится к соответствующей производной. Более строго, в силу теоремы Лагранжа о конечных приращениях и изометричности пространств $\mathbf{L}_2(\Omega)$ и $\mathbf{L}_2(m)$,

$$\left\| \int \frac{e^{i(t+h)\lambda} - e^{it\lambda}}{h} d\mu(\lambda) - \int i\lambda e^{it\lambda} d\mu(\lambda) \right\|_{\mathbf{L}_2(m)} \leq h \sqrt{\int \lambda^2 dm(\lambda)}$$

8.3. Интегралы, производные

и стремится к 0 при уменьшении h в предположении

$$\int \lambda^2 dm(\lambda) < \infty. \quad (8.5)$$

Нетрудно понять, что условие (14.2) является необходимым и достаточным для существования соответствующего предела в среднем квадратическом, а следовательно, и дифференцируемости $\xi(t)$ в изучаемом смысле. Заметим, что для случайных последовательностей (14.2) выполнено в силу конечности меры m и ограниченности области интегрирования.

Окончательно,

$$\frac{d\xi}{dt}(t) = \int i\lambda e^{it\lambda} d\mu(\lambda).$$

Эта формула показывает, что для получения спектрального представления производной стационарного процесса достаточно формально произвести дифференцирование спектрального представления первоначального процесса под знаком интеграла по t . При этом спектральная мера производной процесса будет определяться формулой

$$\mu'(A) = \int_A i\lambda d\mu(\lambda) \Rightarrow d\mu'(\lambda) = i\lambda d\mu(\lambda).$$

Для нахождения спектральной функции производной процесса в среднем квадратическом воспользуемся (7.5). Для ковариационной функции этой производной можно записать

$$K_{d\xi/dt}(t) = \mathbf{M} \left(\frac{d\xi}{dt}(t) \cdot \overline{\frac{d\xi}{dt}(0)} \right) = \int |\lambda|^2 e^{i\lambda t} dm(\lambda),$$

где m – спектральная мера процесса $\xi(t)$, а значит, нужная нам функция находится так:

$$F_{d\xi/dt}(\lambda) = \int_{-\infty}^{\lambda} |s|^2 dm(s).$$

Между прочим, из выписанной формулы немедленно следует, что ковариационная функция рассматриваемого случайного процесса может быть определена лишь в том случае, когда справедливо (14.2). О его справедливости для стационарных последовательностей уже говорилось.

Теперь перейдем к рассмотрению интегральных преобразований. Предположим, что для некоторого интегрируемого с квадратом относительно t «ядра» B

$$\eta(t) = A\xi(t) \equiv \int_T B(t-s)\xi(s)ds.$$

Условимся здесь и ниже в этом разделе считать, что, если T есть множество целых или натуральных чисел, то знак интеграла по T понимается как соответствующая сумма. Тогда во всех случаях – и для непрерывного времени, и для дискретного,

$$\eta(t) = \int_T \left(\int B(t-s)e^{is\lambda} d\mu(\lambda) \right) ds = \int g(\lambda)e^{i\lambda t} d\mu(\lambda),$$

где

$$g(\lambda) = \int_T B(-u)e^{i\lambda u} du.$$

Таким образом, нами фактически доказана следующая

Лемма 5. *Спектральное представление $\eta = A\xi(t)$, где A – оператор описанного выше вида, задается формулой*

$$\eta(t) = \int g(\lambda)e^{i\lambda t} d\mu(\lambda),$$

где функция $g(\lambda)$ строится следующим образом: применим оператор A к функции $e^{it\lambda}$ и подставим в результат $t = 0$:

$$g(\lambda) = Ae^{it\lambda} \Big|_{t=0}$$

Отсюда, в частности, следует, что если F – спектральная функция процесса $\xi(t)$, то для ковариационной функции K_η имеем представление

$$K_\eta(t) = \int e^{i\lambda t} |g(\lambda)|^2 dF(\lambda),$$

а если f_ξ , f_η – соответствующие спектральные плотности, то

$$(\forall \lambda) f_\eta(\lambda) = |g(\lambda)|^2 \cdot f_\xi(\lambda).$$

Продемонстрируем, как эта лемма может быть использована для решения стационарных стохастических дифференциальных уравнений.

8.3. Интегралы, производные

Пусть P – многочлен. Рассмотрим дифференциальное уравнение, формально записываемое в виде

$$P\left(\frac{d}{dt}\right)\eta(t) = \xi(t).$$

Тогда из леммы вытекает, что решение его имеет вид

$$\eta(t) = \int \frac{e^{i\lambda t}}{P(i\lambda)} d\mu(\lambda). \quad (8.6)$$

При этом, аналогично (14.2), должно выполняться

$$\int \frac{dm(\lambda)}{|P(i\lambda)|^2} < \infty$$

(m – структурная функция меры μ). Если последнее условие нарушается, то ни одного стационарного решения рассматриваемого уравнения не существует. Это следует из леммы 2. Проверим, что решение задается формулой (8.6). Действительно, если учесть легко проверяемое индукцией по степени P соотношение

$$P\left(\frac{d}{dt}\right)e^{it\lambda}\Big|_{t=0} = P(i\lambda),$$

то

$$P\left(\frac{d}{dt}\right)\eta(t) = \int \left(P\left(\frac{d}{dt}\right)e^{i\lambda t}\right)\Big|_{t=0} \cdot e^{i\lambda t} \frac{d\mu(\lambda)}{P(i\lambda)} = \xi(t).$$

В частности, стационарное решение уравнения

$$\frac{d\eta}{dt}(t) = \xi(t)$$

существует в том и только том случае, когда

$$\int \frac{dm(\lambda)}{\lambda^2} < \infty,$$

и при этом оно задается формулой

$$\eta(t) = \int \frac{e^{it\lambda}}{i\lambda} d\mu(\lambda) + \zeta,$$

где ζ – произвольная случайная величина, не коррелированная с μ (или, что то же самое, с $Z(\lambda)$).

Если, например, $\xi(t) = A \cos(\eta t + \varphi)$ – процесс из последнего примера, то решение (8.6) таково

$$\eta(t) = \frac{1}{2}A \left(e^{-i\varphi} \frac{e^{-i\eta t}}{-i\eta} + e^{i\varphi} \frac{e^{i\eta t}}{i\eta} \right) + \zeta = \frac{A}{\eta} \sin(\eta t + \varphi) + \zeta.$$

Глава 9

Прогнозирование в частных случаях

К сожалению, общая задача прогнозирования случайных процессов очень сложна, и оптимальное решение ее может быть построено крайне редко. Ниже мы рассмотрим только в некотором смысле «крайние» ситуации, которые допускают относительно несложное решение. В остальных случаях здесь может быть принят тот же порядок действий, как, например, при построении доверительных интервалов, когда изучаемое распределение не является нормальным: мы пользуемся выведенными в частном (нормальном) случае формулами, но заранее знаем, что результат может получиться весьма неточным. Для некоторых процессов, конечно же, постоянно разрабатываются и специфические для них методы прогноза.

9.1 Постановка задачи прогноза

Начнем с постановки задачи прогнозирования в максимальной общности. Пусть $(\forall t) \xi(t) \in \mathbf{L}_2(< \Omega, \mathcal{F}, \mathbf{P} >)$. Требуется, наблюдая значения процесса $\xi(s)$, $s \leq t$, предсказать значение $\xi(q)$ для некоторого известного $q > t$. Обозначим через $\hat{\eta}_{\leq t}$ нужный прогноз по значениям $\xi(s)$, $s \leq t$, наилучший в том смысле, что для выбранного t

$$M(t) = \mathbf{M} |\hat{\eta}_{\leq t} - \xi(q)|^2$$

минимально среди всех допустимых интегрируемых с квадратом прогнозов. Разумеется, вначале следует предположить, что наилучший прогноз существует. Это предположение обычно считается выполненным по умолчанию, по крайней мере, когда допускаются только прогнозы в виде линейных комбинаций величин $\xi(s)$, $s \leq t$. Для такого существования достаточна возможность ортогонального проецирования прогнозируемой величины на линейное подпространство, порожденное случайными величинами $\xi(s)$, $s \leq t$. При этом известно, что среди всех таких комбинаций ортогональная проекция $\xi(q)$ на это пространство наименее удалена от этой величины в смысле нормы пространства $\mathbf{L}_2(\langle \Omega, \mathcal{F}, \mathbf{P} \rangle)$. А это, как раз, и значит, что значение $M(t)$ для такой проекции минимально.

Процесс $\xi(t)$ называется *регулярным слева*, если

$$(\forall q) \lim_{t \rightarrow -\infty} M(t) = \mathbf{D}\xi(q),$$

и *сингулярным слева*, если этот предел равен нулю.

Смысл этих определений состоит в том, что в сингулярном случае любые значения случайного процесса предсказываются абсолютно точно по сколь угодно удаленным по времени в прошлое результатам наблюдений, а в регулярном случае невозможно указать значение истинную величину прогнозируемого значения с точностью более высокой, чем дисперсия этого значения. Как уже говоривалось, рассматриваются прогнозы только стационарных случайных процессов, а следовательно, величина этой дисперсии есть некоторая константа, на зависящая от момента времени q .

Далее, как уже было сказано, мы ограничимся рассмотрением только основных несложных ситуаций в задаче прогнозирования. Прежде всего договоримся далее рассматривать только линейные прогнозы. Это ограничение даст нам возможность придавать задаче прогнозирования простой геометрический смысл поиска ортогональных проекций на линейные подпространства исходного пространства \mathbf{L}_2 . Более того, в определении критерия качества прогноза $M(t)$ мы также ограничимся только поиском его минимума по линейным прогнозам. В таком случае обычно принято определить только что классы случайных процессов называть *линейно регулярными* и *линейно сингулярными*, но мы будем опускать ниже прилагательное «линейный».

Более того, далее относительно подробно рассмотрим только задачи прогноза регулярного и сингулярного процессов. Хотя подоб-

9.2. Сингулярный случай

ные процессы фактически можно считать основными в силу следующего результата.

Теорема 13. Пусть $\mathbf{M}\xi(t) = 0$. Тогда стационарный в широком смысле случайный процесс допускает представление

$$(\forall t) \xi(t) = \text{reg}(t) + \text{sing}(t),$$

где $\text{reg}(t), \text{sing}(t)$ – регулярный и сингулярный процессы соответственно, причем

$$(\forall t) \mathbf{M} \left(\text{reg}(t) \overline{\text{sing}(t)} \right) = 0.$$

Эти процессы определяются однозначно, если потребовать, чтобы они при каждом t были бы измеримы относительно $\xi(s), s \leq t$.

Доказательство этой теоремы можно найти в [11, с. 211]. Пусть теперь $\xi(t), t \in \mathbf{Z}$ – стационарная случайная последовательность. Задачу поиска наилучшего линейного прогноза $\eta = \xi(m)$ по значениям $\xi(t), t \leq 0$ обычно называют *задачей прогноза на m шагов вперед*. Именно ей мы далее и будем заниматься.

Пусть \tilde{m} – спектральная мера последовательности $\xi(t)$. Поскольку $\mathbf{L}_2(\tilde{m})$ изометрично \mathbf{H}_ξ – изометрия осуществляется отображением

$$I(f) = \int_{-\pi}^{\pi} f d\mu, -$$

то для осуществления наилучшего линейного прогноза надо найти такое g в \mathcal{L} – линейной оболочке функций $e^{iz\lambda}, z \leq 0$, чтобы

$$(e^{im\lambda} - g) \perp \mathcal{L},$$

во введенном скалярном произведении т.е.

$$\int_{-\pi}^{\pi} (e^{im\lambda} - g(\lambda)) e^{-in\lambda} d\tilde{m}(\lambda) = 0, \quad n \leq 0. \quad (9.1)$$

9.2 Сингулярный случай

Предположим для начала, что рассматривается задача прогноза на m шагов вперед для сингулярной последовательности $\xi(t)$.

Линейная сингулярность означает возможность аппроксимировать $e^{im\lambda}$, $m > 0$ линейными комбинациями системы функций $e^{in\lambda}$, $n \leq 0$ сколь угодно точно. Это означает, что всегда существует $g(\lambda)$, для которого выполнено (9.1). Рассмотрим несложные достаточные условия сингулярности.

Теорема 14. *Замкнем интервал $[-\pi, \pi)$ в окружность. Если носитель спектральной меры $\xi(t)$ сосредоточен на дуге, занимающей менее половины этой окружности, то последовательность $\xi(t)$ линейно сингулярна.*

Доказательство. Обозначим через z число $e^{-i\lambda}$. Пусть z_0 – середина той дуги окружности, на которой сосредоточен носитель \tilde{m} . Ясно, что заведомо $z_0 \neq 0$. Тогда (это можно легко проверить, сделав соответствующий рисунок) найдется такое действительное число N , что для любого элемента z носителя спектральной меры справедливо $|Nz_0 - z| \leq N$. Записав

$$e^{i\lambda} = \frac{1}{Nz_0 + (z - Nz_0)} = \sum_{n=0}^{\infty} \frac{(Nz_0 - e^{-i\lambda})^n}{(Nz_0)^{n+1}},$$

и применив бином Ньютона, извлекаем отсюда требуемое разложение по отрицательным степеням $e^{i\lambda}$:

$$e^{i\lambda} = \sum_{n=0}^{\infty} \sum_{k=0}^n (-1)^k \frac{C_n^k}{(Nz_0)^{k+1}} e^{-ik\lambda}.$$

Для получения точного прогноза $\xi(m)$ обе части последней формулы нужно возвести в степень m и каждое из вхождений $e^{-ik\lambda}$ в полученной формуле заменить на $\xi(-k)$. Теорема доказана.

Отметим, что мы заодно получили формулы наилучшего линейного прогноза. В частности, формула для прогноза на один шаг выглядит следующим образом:

$$\hat{\xi}(1)_{\leq 0} = \sum_{n=0}^{\infty} \sum_{k=0}^n (-1)^k C_n^k (Nz_0)^{-k-1} \xi(-k).$$

9.3 Прогноз в регулярном случае

9.3.1 Построение прогностических формул

На этот раз пусть $\xi(t)$, $t \in \mathbf{Z}$ – регулярная стационарная случайная последовательность. Поскольку речь опять может идти толь-

9.3. Прогноз в регулярном случае

ко о прогнозе в естественных (и достаточно жестких) ограничениях, условимся предполагать, что у нее определена спектральная плотность. Поскольку регулярность исключает возможность того, что спектральная плотность равна нулю на большей части отрезка $[-\pi, \pi]$ (теорема 14), то, в основном, эта плотность должна быть положительна. Ограничимся здесь ситуацией, когда она положительна на всем отрезке, более того, имеются такие положительные постоянные c_1, c_2 , что для произвольного $\lambda \in [-\pi, \pi]$ выполнено $c_1 < f(\lambda) < c_2$. При сделанных предположениях оказывается, что

$$\mathbf{L}_2(\tilde{m}) = \mathbf{L}_2[-\pi, \pi].$$

Пусть $\mathbf{H}_{\leq 0}$ – замыкание \mathcal{L} в $\mathbf{L}_2[-\pi, \pi]$, $\mathbf{H}_{> 0}$ – замыкание линейной оболочки $e^{in\lambda}$, $n > 0$ в том же пространстве. Из существования и единственности разложения интегрируемых с квадратом функций в ряд Фурье следует, что

$$(\forall g \in \mathbf{L}_2[-\pi, \pi]) (\exists! h_1 \in \mathbf{H}_{\leq 0}, h_2 \in \mathbf{H}_{> 0}) g = h_1 + h_2.$$

Например, условие принадлежности функции пространству $\mathbf{H}_{> 0}$, очевидно, состоит в том, что коэффициенты Фурье с неположительными индексами все равны 0.

Перепишем формулу (9.1) в следующем виде

$$\int_{-\pi}^{\pi} (e^{im\lambda} - g(\lambda)) e^{-in\lambda} f(\lambda) d\lambda = 0, \quad n \leq 0.$$

Тем самым, необходимо найти такую функцию $g \in \mathbf{H}_{\leq 0}$, чтобы

$$(e^{im\lambda} - g(\lambda)) f(\lambda) \in \mathbf{H}_{> 0}. \quad (9.2)$$

Введем еще несколько обозначений. Через $\mathbf{C}_{\geq 0}$ обозначим замыкание линейной оболочки $e^{in\lambda}$, $n \geq 0$ в смысле равномерной сходимости, через $\mathbf{C}_{\leq 0}$ – аналогичное замыкание $e^{in\lambda}$, $n \leq 0$. Так как сходимость в равномерной норме влечет сходимость в \mathbf{L}_2 , то $\mathbf{C}_{\geq 0}, \mathbf{C}_{\leq 0}$ содержатся в $\mathbf{H}_{> 0}, \mathbf{H}_{\leq 0}$ соответственно. Нетрудно понять, что введенные сейчас классы функций являются алгебрами функций, т.е. замкнуты относительно суммирования и умножения. Более того, дополнительно ясно, что они замкнуты и относительно взятия экспонент, т.е. если, например, $g \in \mathbf{C}_{\leq 0}$, то $e^g \in \mathbf{C}_{\leq 0}$ (разложите экспоненту в ряд – он сходится в равномерной норме).

Читателю в качестве несложного упражнения предлагается доказать следующее утверждение.

Лемма 6. *Справедливы следующие утверждения:*

$$\begin{aligned}(g_1 \in \mathbf{H}_{\leq 0}, g_2 \in \mathbf{C}_{\leq 0}) &\implies g_1 g_2 \in \mathbf{H}_{\leq 0}, \\(g_1 \in \mathbf{H}_{> 0}, g_2 \in \mathbf{C}_{\geq 0}) &\implies g_1 g_2 \in \mathbf{H}_{> 0}.\end{aligned}$$

Пусть нашлись такие $f_1 \in \mathbf{C}_{\leq 0}$, $f_2 \in \mathbf{C}_{\geq 0}$, что

$$(\forall \lambda) f(\lambda) = f_1(\lambda) f_2(\lambda).$$

Это представление мы условимся называть *факторизацией спектральной плотности*. В силу единственности разложения в ряд Фурье и того факта, что $f(\lambda)$ – действительное число для произвольного λ , можно (возможно, лишь позаботившись о константе – см. примеры ниже) считать, что $f_1 = \overline{f_2}$ (комплексно сопряжены).

В дальнейшем нам понадобится тот факт, что для найденных нами функций, как правило,

$$\frac{1}{f_1} \in \mathbf{C}_{\leq 0}, \quad \frac{1}{f_2} \in \mathbf{C}_{\geq 0}. \quad (9.3)$$

Это так, поскольку, например, при условии $|1 - f_1| < 1$,

$$\frac{1}{f_1} = \frac{1}{1 - (1 - f_1)} = 1 + (1 - f_1) + (1 - f_1)^2 + \dots,$$

а каждое из слагаемых правой части последнего равенства – элемент $\mathbf{C}_{\leq 0}$. Поэтому соотношения (9.3) мы также будем считать справедливыми.

Далее мы докажем, что факторизация при соблюдении сделанных предположений в регулярном случае возможна всегда и укажем некоторые алгоритмы ее получения.

Нам окажется удобной эквивалентная форма условия (9.2). Для ее получения домножим обе его части на $1/f_2$. В результате получим:

$$(e^{im\lambda} - g(\lambda)) f_1(\lambda) \equiv h(\lambda) \in \mathbf{H}_{> 0},$$

откуда вытекает, что для осуществления прогноза нам требуется представление

$$e^{im\lambda} f_1(\lambda) = g(\lambda) f_1(\lambda) + h(\lambda), \quad g f_1 \in \mathbf{H}_{\leq 0}, h \in \mathbf{H}_{> 0}.$$

Это и есть нужная нам форма условия (9.2).

9.3. Прогноз в регулярном случае

Для построения прогностических формул используем разложение функции f_1 в ряд Фурье. Пусть

$$f_1(\lambda) = c_0 + \sum_{j=1}^{\infty} c_{-j} e^{-ij\lambda}.$$

Отсюда просто выводится

$$e^{im\lambda} f_1(\lambda) = \left(\sum_{k=0}^{m-1} c_{-k} e^{i(m-k)\lambda} + c_{-m} \right) + \sum_{j=1}^{\infty} c_{-m-j} e^{-ij\lambda}.$$

Учитывая, что первая группа слагаемых здесь имеет неотрицательные коэффициенты при мнимой единице в показателях экспонент, а вторая отрицательные, можно выбрать

$$g(\lambda) = \frac{c_{-m} + \sum_{j=1}^{\infty} c_{-m-j} e^{-ij\lambda}}{f_1(\lambda)}.$$

Прибегнем к разложению Фурье еще раз, на этот раз разложим полученную функцию g . Если

$$g(\lambda) = b_0 + \sum_{j=1}^{\infty} b_{-j} e^{-ij\lambda},$$

то, как следует из изометричности соответствующих пространств, наилучший линейный прогноз на m шагов задается следующим образом:

$$\hat{\xi}_{\leq 0} = \sum_{j=0}^{\infty} b_{-j} \xi(-j).$$

Вычислим среднеквадратическую ошибку полученного только что прогноза на m шагов, вновь воспользовавшись изометричностью пространств $\mathbf{L}_2(\Omega)$ и $\mathbf{L}_2(\tilde{m})$.

$$\sigma^2(m) = \int_{-\pi}^{\pi} |e^{im\lambda} - g(\lambda)|^2 f(\lambda) d\lambda.$$

В силу сделанных выше допущений $f(\lambda) = f_1(\lambda) \overline{f_1(\lambda)}$, а значит,

$$\sigma^2(m) = \int_{-\pi}^{\pi} |(e^{im\lambda} - g(\lambda)) f_1(\lambda)|^2 d\lambda = \int_{-\pi}^{\pi} |h(\lambda)|^2 d\lambda.$$

Отсюда, с учетом ортонормированности в пространстве \mathbf{L}_2 базиса $\frac{1}{\sqrt{2\pi}}e^{in\lambda}$, $n \in \mathbf{Z}$, получаем

$$\sigma^2(m) = \int_{-\pi}^{\pi} \left| \sum_{j=0}^{m-1} c_j e^{i(m-j)\lambda} \right|^2 d\lambda = 2\pi \sum_{k=1}^m |c_{-m+k}|^2.$$

В силу свойства стационарности рассмотрение прогноза на бесконечно большое число шагов совпадает с задачей прогнозирования по «предельно далекому прошлому». При $m \rightarrow \infty$ видим, что ошибка такого прогноза

$$\sigma^2(m) \rightarrow 2\pi \sum_{j=0}^{\infty} |c_{-j}|^2 = \int_{-\pi}^{\pi} |f_1(\lambda)|^2 d\lambda = K(0),$$

что совпадает с дисперсией случайной величины $\xi(m)$, чем подтверждается наличие здесь регулярного случая.

Итак, для построения оптимальных прогностических формул нам, кроме стандартных навыков разложения функции в ряд Фурье, потребуются умение находить или оценивать спектральную плотность, а также осуществлять ее факторизацию. Первая из этих задач скорее практическая и относится к разделу статистики случайных процессов. А вот решить вторую можно «теоретически». Поэтому начнем со второй задачи.

9.3.2 Два способа факторизации плотности

Сначала – общие рекомендации. Пусть спектральная плотность f достаточно гладкая. Предположив, что логарифм спектральной плотности допускает разложение в равномерно сходящийся ряд Фурье, рассмотрим это разложение.

$$\ln f(\lambda) = \sum_{z \in \mathbf{Z}} a_z e^{iz\lambda}.$$

При всех значениях аргумента этот логарифм является действительным числом, т.к. $f(\lambda)$ в изучаемом случае принимает только строго положительные значения. Отсюда

$$a_0 \in \mathbf{R}, \quad (\forall z) a_{-z} = \overline{a_z}.$$

9.3. Прогноз в регулярном случае

Теперь можно выбрать

$$R_1(\lambda) = \frac{a_0}{2} + \sum_{z<0} a_z e^{iz\lambda}, \quad R_2(\lambda) = \frac{a_0}{2} + \sum_{z>0} a_z e^{iz\lambda}.$$

Тогда $R_1 \in \mathbf{C}_{\leq 0}$, $R_2 \in \mathbf{C}_{\geq 0}$, $\overline{R_1} = R_2$. Положим

$$f_1(\lambda) = \exp\{R_1(\lambda)\}, \quad f_2(\lambda) = \exp\{R_2(\lambda)\}.$$

Нетрудно понять, что эти функции и осуществляют нужную факторизацию (см. (9.3)). При этом, как мы и предполагали выше, $f_2 = \overline{f_1}$ в силу только что отмеченных соотношений между коэффициентами Фурье в разложениях R_1, R_2 .

Получим здесь формулу для среднеквадратической ошибки прогноза. Используя разложение экспоненты в степенной ряд, можно определить коэффициенты разложения Фурье функции f_1 через соответствующие коэффициенты R_1 :

$$f_1(\lambda) = 1 + \sum_{j=1}^{\infty} \frac{R_1^j(\lambda)}{j!} = 1 + \sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{a_0}{2} + \sum_{z<0} a_z e^{iz\lambda} \right)^j.$$

В частности, отсюда

$$c_0 = 1 + \sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{a_0}{2} \right)^j = e^{a_0/2},$$

а следовательно, среднеквадратическая ошибка прогноза за один шаг

$$\sigma^2(1) = 2\pi |c_0|^2 = 2\pi e^{a_0},$$

в результате чего получаем красивую формулу

$$\sigma^2(1) = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(\lambda) d\lambda \right\}.$$

На практике иногда можно обойтись и без разложений Фурье. Рассмотрим важный частный случай, когда спектральная плотность может быть представлена отношением двух многочленов, т.е.

$$f(\lambda) = \frac{P(e^{i\lambda})}{Q(e^{i\lambda})},$$

где P, Q – многочлены. Излагаемый ниже метод известен под названием метода Яглома. Если комплексное число z лежит на единичной окружности с центром в начале координат, то из нашего представления следует, что $P(z)/Q(z)$ – действительное число, а значит каждому корню z_i одного из многочленов, отличному от 0, соответствует корень $(\bar{z}_i)^{-1}$, симметричный z_i относительно единичной окружности с центром в начале координат. Это вытекает из так называемого принципа симметрии Римана – Шварца (см., например, [8, с. 158 - 162]).

Тогда наша дробь может быть записана в виде

$$\frac{P(z)}{Q(z)} = c' z^k \frac{(z - z_1)(z - \bar{z}_1^{-1}) \dots (z - z_n)(z - \bar{z}_n^{-1})}{(z - w_1)(z - \bar{w}_1^{-1}) \dots (z - w_m)(z - \bar{w}_m^{-1})},$$

где c' – комплексная постоянная, k – разность кратностей корня 0 у многочленов P и Q . Условимся считать, что корни z_i, w_j по модулю больше 1, а сопряженные их обратным, соответственно, меньше 1. Заметим, что

$$(z - z_i)(z - \bar{z}_i^{-1}) = \frac{(z_i - z)(\bar{z}_i - z^{-1})}{-z^{-1}\bar{z}_i}.$$

Проделав это для каждой пары соответственных корней обоих многочленов, получим

$$\frac{P(z)}{Q(z)} = c z^{k-n+m} \frac{(z_1 - z)(\bar{z}_1 - z^{-1}) \dots (z_n - z)(\bar{z}_n - z^{-1})}{(w_1 - z)(\bar{w}_1 - z^{-1}) \dots (w_m - z)(\bar{w}_m - z^{-1})}.$$

Если $z = e^{i\lambda}$, то $z^{-1} = \bar{z}$ и

$$(z_j - z)(\bar{z}_j - z^{-1}) = |z_j - z|^2 \geq 0.$$

Следовательно, c – действительная постоянная, $k + n - m = 0$ и можно выбрать

$$f_1(\lambda) = \sqrt{c} \frac{(\bar{z}_1 - e^{-i\lambda}) \dots (\bar{z}_n - e^{-i\lambda})}{(\bar{w}_1 - e^{-i\lambda}) \dots (\bar{w}_m - e^{-i\lambda})}. \quad (9.4)$$

Надо лишь убедиться, что $f_1, f_1^{-1} \in \mathbf{C}_{\leq 0}$. Для этого достаточно показать, что для произвольного комплексного a , по модулю большего 1, справедливо $(a - e^{i\lambda})^{-1} \in \mathbf{C}_{\leq 0}$. Запишем

$$\frac{1}{a - e^{-i\lambda}} = \frac{\frac{1}{a}}{1 - \frac{e^{-i\lambda}}{a}} = \frac{1}{a} \left(1 + \sum_{j=1}^{\infty} \left(\frac{e^{-i\lambda}}{a} \right)^j \right).$$

9.3. Прогноз в регулярном случае

Последняя сумма очевидно лежит в $\mathbf{C}_{\leq 0}$. При выводе мы воспользовались разложением Маклорена и тем фактом, что в силу выбора большого по модулю a

$$\left| \frac{e^{-i\lambda}}{a} \right| < 1.$$

Подведем итог. Для получения факторизации спектральной плотности в частном случае достаточно разложить числитель и знаменатель дроби на линейные множители, выбрать корни, по модулю большие 1, задать f_1 по формуле (9.4), а $f_2 = f/f_1$.

9.3.3 Несколько несложных примеров

Зафиксируем действительное число s . Как будет ясно из дальнейшего, можно без ограничения общности считать, что $s \in [-\pi, \pi]$. Выберем некоторое комплексное число g . Пусть случайная величина ξ такова, что $\mathbf{M}\xi = 0$, $\mathbf{M}\xi^2 = 1$. Введем в рассмотрение случайную последовательность

$$\xi_t = g\xi e^{ist}, \quad t \in \mathbf{Z}.$$

Нетрудно проверить, эта случайная последовательность стационарна, причем

$$K(n) = |g|^2 e^{isn}, \quad n \in \mathbf{Z} \Rightarrow K(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dm(\lambda),$$

где мера m сосредоточена в точке s и равна в этой точке $|g|^2$. Стохастическая спектральная мера последовательности, также сосредоточенная в точке s , равна в ней $g\xi$. Итак, мы имеем дело с сингулярным случаем, и в формулах прогноза раздела 5.4 можно взять $z_0 = e^{is}$, $N = 1$.

Слегка обобщим рассмотренный пример. Пусть комплекснозначные случайные элементы z_1, \dots, z_N таковы, что

$$\mathbf{M}z_j = 0, \quad \mathbf{M}|z_j|^2 = \sigma_j^2, \quad \mathbf{M}z_j \bar{z}_k = 0$$

при всех j, k и $k \neq j$. Фиксируем числа s_k , $k = 1, \dots, N$ в интервале $[-\pi, \pi]$ и определим стационарную случайную последовательность

$$\xi_n = \sum_{k=1}^N z_k e^{is_k n}, \quad n \in \mathbf{Z}.$$

Понятно, что для соответствующей ковариационной функции

$$K(n) = \sum_{k=1}^N \sigma_k^2 e^{i s_k n}$$

спектральная мера сосредоточена в точках s_1, \dots, s_N и принимает в них значения $\sigma_1^2, \dots, \sigma_N^2$. Вопрос о наличии или отсутствии сингулярности здесь решается на основе расположения этих точек в интервале $[-\pi, \pi]$.

В этих примерах можно говорить о том, что соответствующие ковариационные функции являются комбинацией небольшого числа гармоник (функций $e^{i s_k n}$, $k = 1, \dots, N$). Следующий пример показывает ситуацию, когда в равной степени задействованы все гармоники. Такой процесс принято называть белым шумом по аналогии с белым цветом, в котором равным образом участвуют все цвета спектра.

Пусть случайные элементы ξ_k , $k \in \mathbf{Z}$ независимы, одинаково распределены, имеют нулевые математические ожидания, и $\mathbf{M}|\xi_k|^2 = 1$, $k \in \mathbf{Z}$. Тогда стационарную случайную последовательность ξ_n , $n \in \mathbf{Z}$ мы назовем *белым шумом*. Ее ковариационная функция $K(n)$, равная 1 при $n = 0$ и нулю при остальных n , представима в виде

$$K(n) = \int_{-\pi}^{\pi} e^{i \lambda n} dF(\lambda),$$

где F – функция равномерного распределения на $[-\pi, \pi]$. В обозначениях предыдущего раздела,

$$f(\lambda) = \frac{1}{2\pi}, \quad f_1(\lambda) = \frac{1}{\sqrt{2\pi}}, \quad g(\lambda) = 0,$$

в силу чего мы имеем дело с регулярным случаем, но наилучший прогноз здесь равен 0, и величина ошибки даже за один шаг совпадает с дисперсией процесса.

Естественно, что более содержательные примеры регулярного случая связаны с более сложно устроенными спектральными плотностями.

Пусть, например, спектральная плотность задается формулой

$$f(\lambda) = 5 + 4 \cos \lambda.$$

Тогда

$$f(\lambda) = 5 + 2e^{i\lambda} + 2e^{-i\lambda}.$$

9.3. Прогноз в регулярном случае

Заметим, что ковариационная функция соответствующей стационарной последовательности имеет вид

$$K(t) = \int_{-\pi}^{\pi} e^{i\lambda t} f(\lambda) d\lambda = \frac{10 \sin \pi t}{t} + \frac{4 \sin \pi(t+1)}{t+1} + \frac{4 \sin \pi(t-1)}{t-1}.$$

При всех t отличных от 0 и ± 1 это выражение равно 0, а в перечисленных точках формально не определено. Но по непрерывности естественно можно доопределить его, приняв за значения функции $K(0) = 10\pi$, $K(\pm 1) = 4\pi$. Запишем

$$f(\lambda) = \frac{2e^{2i\lambda} + 5e^{i\lambda} + 2}{e^{i\lambda}} = \frac{P(e^{i\lambda})}{Q(e^{i\lambda})}.$$

Разлагая числитель на множители, получим

$$f(\lambda) = 2 \frac{(e^{i\lambda} + \frac{1}{2})(e^{i\lambda} + 2)}{e^{i\lambda}} = (2 + e^{-i\lambda})(e^{i\lambda} + 2).$$

Это означает, что можно выбрать

$$f_1(\lambda) = 2 + e^{-i\lambda}, \quad f_2(\lambda) = e^{i\lambda} + 2,$$

и, в обозначениях предыдущего подраздела, $c_0 = 2$, $c_{-1} = 1$, $c_{-k} = 0$ при $k \geq 4$.

Выишем формулу прогноза на 1 шаг ($m = 1$).

$$g(\lambda) = c_{-1} f_1^{-1}(\lambda) = \frac{1}{2} \frac{1}{1 + \frac{1}{2} e^{-i\lambda}},$$

следовательно, применяя формулу суммы бесконечно убывающей геометрической прогрессии, получим:

$$g(\lambda) = \frac{1}{2} \left(1 + \sum_{j=1}^{\infty} \frac{(-1)^j e^{-ij\lambda}}{2^j} \right).$$

Окончательно, наилучший линейный прогноз на один шаг задается формулой

$$\hat{\xi}(1)_{\leq 0} = \frac{1}{2} \xi(0) + \sum_{j=1}^{\infty} (-1)^j \frac{\xi(-j)}{2^{j+1}}.$$

Вычислим величину ошибки прогноза:

$$\sigma^2(1) = 2\pi c_0^2 = 8\pi,$$

что указывает на низкую надежность прогноза. Если бы мы взяли прогнозировать на 2 шага, то получили бы, что

$$\sigma^2(2) = 2\pi(c_0^2 + c_{-1}^2) = 10\pi = \mathbf{D}\xi(0),$$

что подтверждает регулярность рассматриваемой последовательности – невозможно дать прогноз точнее, чем с точностью до дисперсии. Таким образом, при прогнозе на два и более шагов наилучшим прогнозом следует признать математическое ожидание $a = \mathbf{M}\xi(t)$ (это константа, поскольку мы работаем только со стационарными процессами), или его оценку

$$a^* = \frac{\xi(-N) + \dots + \xi(0)}{N + 1}.$$

9.4 Задача статистического оценивания спектральной плотности

Итак, знание спектральной плотности наблюдаемого случайного процесса является крайне важным для решения задач прогноза. С помощью спектральной плотности решаются также многие другие интересные задачи (например, с ее помощью выписываются решения стационарных дифференциальных уравнений, о чем уже говорилось выше). Сейчас мы, видимо, впервые в настоящем курсе, рассмотрим практический подход к задаче определения спектральной плотности. Поскольку в практических ситуациях мы имеем возможность воспользоваться только отдельными числовыми значениями, наблюдениями за развитием конкретного случайного процесса, то речь здесь может идти только об оценке нужной нам функции. Таким образом, мы имеем дело со специальной статистической задачей оценивания. Но, с другой стороны, мы по числовым значениям пытаемся восстановить функцию, т.е. имеем дело с задачей интерполяции. Это обычная ситуация при работе со статистикой случайных процессов.

Наиболее важным для нас, как понятно из предыдущего подраздела, является случай стационарного процесса. А поскольку непрерывно наблюдать за процессом невозможно, будем без ограничения общности считать, что мы изучаем стационарную случайную

9.4. Оценка спектральной плотности

последовательность. Обозначим постоянное в этом случае математическое ожидание $\mathbf{M}\xi_n = m$, $n \in Z$. Пусть также

$$K(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda) -$$

соответствующая теоретическая ковариационная функция. В нашем распоряжении имеются N наблюдений элементов случайной последовательности x_0, x_1, \dots, x_{N-1} . Обычной оценкой m является, конечно же, их среднее арифметическое m^* . Для упрощения формул будем считать, что эта, или какая-то другая состоятельная несмещенная оценка среднего вычтена из всех наблюдаемых значений, т.е. их среднее равно 0. Для m любого знака число тех пар (i, j) , для которых разность $i - j = m$, есть $N - |m|$. Тогда естественной оценкой для функции K является

$$\hat{K}_N(n) = \begin{cases} \frac{1}{N-n} \sum_{k=0}^{N-n-1} x_{k+n} \bar{x}_k, & n \geq 0; \\ \frac{1}{N+n} \sum_{k=-n}^N x_{k+n} \bar{x}_k, & n < 0. \end{cases}$$

Из стандартных результатов математической статистики следует, что $\hat{K}_N(n)$ является несмещенной оценкой $K(n)$.

Ориентируясь на формулы обратного преобразования Фурье, введем в рассмотрение

$$f_N(\lambda) = \frac{1}{2\pi N} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} K(s-t) e^{-i\lambda(s-t)}.$$

Эта функция неотрицательна в силу (8.1). Ясно, что

$$f_N(\lambda) = \frac{1}{2\pi} \sum_{|m| < N} K(m) \left(1 - \frac{|m|}{N}\right) e^{-i\lambda m}.$$

Если $F(\lambda)$ – функция распределения с плотностью f_N , то

$$\int_{-\pi}^{\pi} e^{i\lambda n} dF_N(\lambda) = \begin{cases} \left(1 - \frac{|n|}{N}\right) K(n), & |n| < N, \\ 0, & |n| \geq N \end{cases}$$

в силу ортонормированности системы функций $\frac{1}{\sqrt{2\pi}}e^{-i\lambda n}$ при $n \in \mathbf{Z}$.
Пусть

$$F(\lambda) = \lim_{N \rightarrow \infty} F_N(\lambda)$$

в смысле слабой сходимости. Тогда из выписанных формул следует, что F – спектральная функция, а, соответственно, $f(\lambda) = \lim_{N \rightarrow \infty} f_N(\lambda)$ – спектральная плотность нашей последовательности.

Заменим K ее оценкой. Статистика

$$\hat{f}_N(\lambda) = \frac{1}{2\pi} \sum_{|m| < N} \hat{K}_N(m) \left(1 - \frac{|m|}{N}\right) e^{-i\lambda m},$$

являющаяся оценкой $f_N(\lambda)$, называется *периодограммой*. Оказывается, для периодограммы может быть выписана более простая формула

$$\hat{f}_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{s=0}^{N-1} x_s e^{-i\lambda s} \right|^2. \quad (9.5)$$

Действительно, правую часть (9.5) распишем в виде

$$\frac{1}{2\pi N} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} x_s \bar{x}_t e^{-i\lambda(s-t)} = \frac{1}{2\pi N} \sum_{|m| < N} \sum_t x_{t+m} \bar{x}_t e^{-i\lambda m},$$

и, поскольку во внутренней сумме $N - |m|$ слагаемых, то можно продолжить выписываемое равенство до

$$\frac{1}{2\pi N} \sum_{|m| < N} (N - |m|) \hat{K}(m) e^{-i\lambda m} = \hat{f}_N(\lambda).$$

Наша цель – заметить, что можно использовать (9.5) в качестве оценки спектральной плотности. Т.к. $\mathbf{M}\hat{K}(n) = K(n)$, то $\hat{f}_N(\lambda)$ является несмещенной оценкой $f_N(\lambda)$. Заметим далее, что

$$\begin{aligned} f_N(\lambda) &= \frac{1}{2\pi N} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} K(s-t) e^{-i\lambda(s-t)} = \\ &= \frac{1}{2\pi N} \int_{-\pi}^{\pi} \left| \sum_{k=0}^{N-1} e^{i(\mu-\lambda)k} \right|^2 f(\mu) d\mu. \end{aligned}$$

9.4. Оценка спектральной плотности

Функция

$$\Phi_M(s) = \frac{1}{2\pi M} \left| \sum_{k=0}^M e^{isk} \right|^2 = \frac{1}{2\pi M} \left| \frac{\sin(sM/2)}{\sin(s/2)} \right|^2$$

обычно называется *ядром Фейера*. Из курса функционального анализа известна следующая сходимость, связанная с этим ядром:

$$\mathbf{M}\hat{f}_N(\lambda) = f_N(\lambda) = \int_{-\pi}^{\pi} \Phi_{N-1}(\mu - \lambda) f(\mu) d\mu \rightarrow f(\lambda)$$

при любом λ и $N \rightarrow \infty$. Итак, $\hat{f}_N(\lambda)$ является асимптотически несмещенной оценкой для $f(\lambda)$.

Но в практических задачах ошибки двух в целом неплохих приближений обычно накапливаются, что приводит к тому, что оценка (9.5) не может использоваться в сколько-нибудь удовлетворительном смысле. Пусть, например, ξ_n – белый шум, составленный из стандартных нормальных случайных величин. Тогда, как мы уже знаем, $f(\lambda) \equiv \frac{1}{2\pi}$, а

$$\hat{f}_N(\lambda) = \frac{1}{2\pi} \left| \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-i\lambda k} \right|^2.$$

Даже при $\lambda = 0$ у нас $\hat{f}_N(0) = \frac{1}{2\pi} \xi^2$, где ξ имеет стандартное нормальное распределение, и

$$\mathbf{M}|\hat{f}_N(0) - f(0)|^2 = \frac{1}{4\pi^2} \mathbf{M}|\xi^2 - 1|^2 \rightarrow 0.$$

Для преодоления отмеченных явлений оценку (9.5) обычно улучшают за счет введения специальной весовой функции W_N , называемой *спектральным окном*, и полагают

$$f_N^W(\lambda) = \int_{-\pi}^{\pi} W_N(\lambda - \nu) \hat{f}_N(\nu) d\nu.$$

От спектрального окна при этом требуют

1. $W_N(0)$ – максимальное значение W_N ;

2. $\int_{-\pi}^{\pi} W_N(\lambda) d\lambda = 1;$
3. $(\forall \lambda) \lim_{N \rightarrow \infty} \mathbf{M} |\hat{f}_N^W(\lambda) - f(\lambda)|^2 = 0.$

Можно предложить много способов построения спектральных окон. Вот наиболее распространенные. Пусть числовая последовательность $a_N = o(N) \rightarrow \infty$ выбрана произвольно,

$$W_N(\lambda) = a_N B(a_N \lambda).$$

Если

$$B(\mu) = \frac{1}{2\pi} \left| \frac{\sin(\mu/2)}{\mu/2} \right|^2,$$

то мы получаем *окно Бартлетта*. Если в тех же обозначениях выбрать

$$B(\mu) = \frac{3}{8\pi} \left| \frac{\sin(\mu/4)}{\mu/4} \right|^4,$$

то получаем *окно Парзена*. Наконец, последовательность a_N можно брать не произвольной, а построить специальным образом, взяв любое $\alpha \in (0, 2]$ и, если ввести

$$B(\mu) = \begin{cases} \frac{\alpha+1}{2^\alpha} (1 - |\mu|^\alpha), & |\mu| \leq 1, \\ 0, & |\mu| > 1, \end{cases}$$

то мы получим *окно Журбенко*. Каждый из этих подходов имеет свои достоинства и недостатки, которые здесь мы обсуждать не будем.

Глава 10

Процессы размножения и гибели

10.1 Основные предположения

В этом разделе мы обратимся к изучению одного из конкретных видов марковских процессов, который оказывается полезным при моделировании многих практических явлений в области природы, техники и социальных эволюций.

Пусть в некоторой области имеются частицы. Будем считать, что их количество в этой области случайно и обозначим $p_n(t)$ вероятность того, что в момент времени t имелось ровно n частиц. Условимся считать, что возможные количество частиц в области являются состояниями устойчивой цепи Маркова, а изменение этого количества со временем описывается переходными вероятностями такой цепи. Тем самым, сделано предположение, что вероятности перехода удовлетворяют условиям 1 – 4 раздела 6.

Отдельные члены последовательности чисел λ_n из упомянутых условий здесь естественно назвать интенсивностями размножения, а числа μ_n – интенсивностями гибели. Обе эти последовательности считаются заданными заранее и полностью известными. Если оказывается, что $(\forall n) (\mu_n = 0)$, но, хотя бы некоторые из λ_n положительны, то говорят о чистом размножении, если же $(\forall n) (\lambda_n = 0)$ одновременно с наличием положительных μ_n , то о чистой гибели. Заметим, что, как и ранее, из сформулированных допущений следу-

ет, что вероятность рождения или гибели за время Δt одновременно двух или более частиц есть $o(\Delta t)$.

Укажем один набор достаточных условий для того, чтобы были справедливы все сделанные предположения. Пусть частицы размножаются и гибнут независимо друг от друга, причем каждая из них за промежуток Δt порождает новую с вероятностью b или гибнет с вероятностью d ($b + d \leq 1$). Предположим, что

$$b = \lambda\Delta t + o(\Delta t), \quad d = \mu\Delta t + o(\Delta t).$$

При этом очевидно, что вероятность за время Δt родиться или погибнуть двум или большему количеству частиц будет равна $o(\Delta t)$ независимо от того, сколько в области имелось к этому моменту частиц. Теперь, обозначая $b_n(\Delta t)$, $d_n(\Delta t)$ числа родившихся и погибших частиц, получим

$$\mathbf{P}(b_n(\Delta t) = 1) = n\lambda\Delta t + o(\Delta t), \quad \mathbf{P}(d_n(\Delta t) = 1) = n\mu\Delta t + o(\Delta t),$$

а значит,

$$\begin{aligned} P_{n,n+1}(\Delta t) &= \mathbf{P}(b_n(\Delta t) = 1, d_n(\Delta t) = 0) + o(\Delta t) = \\ &= (n\lambda\Delta t + o(\Delta t))(1 - n\mu\Delta t + o(\Delta t)) + o(\Delta t) = n\lambda\Delta t + o(\Delta t); \\ P_{n,n-1}(\Delta t) &= n\mu\Delta t + o(\Delta t); \\ P_{n,n}(\Delta t) &= \mathbf{P}(b_n(\Delta t) = 0, d_n(\Delta t) = 0) + \mathbf{P}(b_n(\Delta t) = 1, d_n(\Delta t) = 1) + \\ &+ o(\Delta t) = 1 - n(\lambda + \mu)\Delta t + o(\Delta t). \end{aligned}$$

Тем самым, в этом случае все сделанные предположения оказываются выполненными при $\lambda_n = n\lambda$, $\mu_n = n\mu$, то есть, согласно терминологии, введенной в разделе 5.1, мы получили устойчивую цепь Маркова специального типа, и для корректного применения к некоторому явлению результатов настоящего раздела вполне хватит проверки упомянутых достаточных условий.

10.2 Одномерные распределения

Далее, если не будет оговорено нечто иное, условимся считать, что процесс размножения и гибели является устойчивой це-

10.2. Одномерные распределения

пью Маркова специального типа (выполнено (6.3)). Получим дифференциальные уравнения, описывающие такой процесс размножения и гибели. По формуле полной вероятности

$$p_n(t + \Delta t) = p_n(t)(1 - n\lambda\Delta t + o(\Delta t))(1 - n\mu\Delta t + o(\Delta t)) + \\ + p_{n-1}(t)((n-1)\lambda\Delta t + o(\Delta t))(1 - (n-1)\mu\Delta t + o(\Delta t)) + \\ + p_{n+1}(t)((n+1)\mu\Delta t + o(\Delta t))(1 - (n+1)\lambda\Delta t + o(\Delta t)) + o(\Delta t).$$

Отсюда нетрудно получить

$$\frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = -n(\lambda + \mu)p_n(t) + (n-1)\lambda p_{n-1}(t) + \\ + (n+1)\mu p_{n+1}(t) + o(1).$$

Переходя к пределу при $\Delta t \rightarrow 0$, имеем

$$p'_n(t) = -n(\lambda + \mu)p_n(t) + (n-1)\lambda p_{n-1}(t) + (n+1)\mu p_{n+1}(t).$$

Выше предполагалось, что $n \geq 1$. Если же $n = 0$, то аналогичными рассуждениями просто выводится

$$p'_0(t) = \mu p_1(t).$$

Заметим, что полученное уравнение при $n = 0$ можно считать частным случаем предыдущего уравнения, если принять во внимание два обстоятельства. Во-первых, $p_0(t) = 0$, поскольку если бы в какой-то момент в области не оказалось частиц, то дальнейшая эволюция процесса стала бы невозможна. Во-вторых, при рассматриваемом n вероятность $p_{n-1}(t)$ нужно принять равной 0 по понятным причинам.

Умножим n -ное уравнение полученной системы на x^n и, складывая, получим

$$\sum_{n=0}^{\infty} p'_n(t)x^n = \mu \sum_{n=1}^{\infty} p_n(t)x^{n-1}n + \lambda x^2 \sum_{n=1}^{\infty} p_n(t)x^{n-1}n - \\ - (\lambda + \mu) \sum_{n=0}^{\infty} p_n(t)x^n n.$$

Если мы теперь обозначим

$$F(x, t) = \sum_{n=0}^{\infty} p_n(t)x^n,$$

то полученное выше уравнение можно переписать в виде

$$\frac{\partial F}{\partial t} = (\mu - (\lambda + \mu)x + \lambda x^2) \frac{\partial F}{\partial x}.$$

Предположим для простоты, что $p_1(0) = 1$, т.е. в начальный момент времени в области находилась одна частица. Отсюда, в частности, следует, что $p_n(0) = 0$ при $n \neq 1$. Чтобы решить уравнение в частных производных, составим обыкновенное дифференциальное уравнение:

$$dt = \frac{-dx}{\mu - (\lambda + \mu)x + \lambda x^2},$$

откуда

$$\begin{aligned} t - c &= \frac{1}{\mu - \lambda} \int \left(\frac{1}{1-x} - \frac{\lambda}{\mu - \lambda x} \right) dx = \\ &= \frac{1}{\mu - \lambda} \ln \left| \frac{\mu - \lambda x}{1-x} \right| - \end{aligned}$$

первый интеграл линейного уравнения в частных производных первого порядка. Тогда, как известно, общим решением уравнения в частных производных будет

$$F(t, x) = R \left(\frac{1}{\mu - \lambda} \ln \left| \frac{\mu - \lambda x}{1-x} \right| + t \right),$$

где R – произвольная дифференцируемая функция. Чтобы найти эту функцию в заявленных начальных (при $t = 0$) условиях, заметим, что при $t = 0$ в силу сделанных допущений $F(x, t) = x$, или, тождественно по x ,

$$R \left(\frac{1}{\mu - \lambda} \ln \left| \frac{\mu - \lambda x}{1-x} \right| \right) = x,$$

и R есть обратная функция для $u = \frac{1}{\mu - \lambda} \ln \left| \frac{\mu - \lambda x}{1-x} \right|$. Следовательно,

$$R(u) = \frac{\mu - e^{(\mu - \lambda)u}}{\lambda - e^{(\mu - \lambda)u}}.$$

10.2. Одномерные распределения

Подставляя в это выражение вместо u аргумент R , имеем

$$F(x, t) = \frac{\mu(1 - e^{(\mu-\lambda)t}) + x(\lambda e^{(\mu-\lambda)t} - \mu)}{\lambda - \mu e^{(\mu-\lambda)t} + \lambda x(e^{(\mu-\lambda)t} - 1)}.$$

Разлагая F в ряд по степеням x , нетрудно получить, что

$$p_n(t) = \frac{(\mu - \lambda)^2 e^{(\mu-\lambda)t} (1 - e^{(\mu-\lambda)t})^{n-1} \lambda^{n-1}}{(\lambda - \mu e^{(\mu-\lambda)t})^{n+1}}, \quad n \geq 1,$$

$$p_0(t) = \frac{\mu(1 - e^{(\mu-\lambda)t})}{\lambda - \mu e^{(\mu-\lambda)t}},$$

А также

$$p_{\geq 1}(t) = 1 - p_0(t) = \sum_{k=1}^{\infty} p_k(t) = \frac{\lambda - \mu}{\lambda - \mu e^{(\mu-\lambda)t}}.$$

Пусть $\mu < \lambda$. Тогда при $t \rightarrow \infty$

$$p_0(t) \rightarrow \frac{\mu}{\lambda}, \quad p_{\geq 1}(t) \rightarrow 1 - \frac{\mu}{\lambda},$$

а если, наоборот, $\mu > \lambda$, то

$$p_0(t) \rightarrow 1, \quad p_{\geq 1}(t) \rightarrow 0.$$

Отсюда можно сделать очевидные выводы. Например, если гибель частиц происходит с большей интенсивностью, то через достаточно продолжительное время частиц в области наверняка не останется.

Заметим, что из рассмотренных случаев выпала ситуация, когда $\lambda = \mu$, т.е. размножение и гибель происходят с равными интенсивностями. Для корректного рассмотрения этой ситуации придется вернуться к решенному выше линейному уравнению в частных производных и заново решить его. Ситуация получится более простой, поскольку ни логарифм, ни экспонента в первом интеграле здесь уже не появятся. Предоставляю читателю самостоятельно проделать эти выкладки и проинтерпретировать полученный результат.

10.3 Примеры

10.3.1 Линейный рост с иммиграцией

Пусть в модели процесса размножения и гибели $\lambda_n = \lambda n + a$, $\mu_n = \mu n$, что можно интерпретировать как наличие постоянной интенсивности роста и гибели при наличии неизменной относительно общего числа частиц составляющей роста a . Эту составляющую удобно интерпретировать как иммиграцию, не связанную с количеством частиц в области. Привлекая систему прямых уравнений Колмогорова, получим

$$\begin{aligned} P'_{n,0}(t) &= -aP_{n,0} + \mu P_{n,1}(t), \\ P'_{n,j}(t) &= (a + \lambda(j-1))P_{n,j-1}(t) - \\ &- ((\lambda + \mu)j + a)P_{n,j}(t) + \mu(j+1)P_{n,j+1}(t), \quad j \geq 1. \end{aligned}$$

Умножим j -е уравнение на j и суммируем. Обозначим $M(t)$ математическое ожидание процесса и допустим, что $\xi(0) = n$ с вероятностью 1. Тогда получим

$$M'(t) = a + (\lambda - \mu)M(t).$$

Если $\lambda \neq \mu$, то, проинтегрировав это уравнение, получим

$$M(t) = Ce^{(\lambda - \mu)t} - \frac{a}{\lambda - \mu}.$$

Подставляя сюда $t = 0$, находим

$$c = n + \frac{a}{\lambda - \mu}.$$

Окончательно,

$$M(t) = \frac{a}{\lambda - \mu} \left(e^{(\lambda - \mu)t} - 1 \right) + ne^{(\lambda - \mu)t}.$$

Отметим, что при $\lambda > \mu$ значение $M(t)$ неограниченно увеличивается с течением времени, при $\lambda < \mu$, монотонно убывая, стремится к $a/(\mu - \lambda)$.

Если же $\lambda = \mu$, то

$$M(t) = at + n, \quad M(t) \rightarrow \infty.$$

10.3. Примеры

Очевидная возможная интерпретация – при равных интенсивностях размножения и гибели популяция продолжает расти исключительно за счет иммиграции. В отличие от экспоненциального ее роста в случае, когда размножение превалировало над гибелью, данный рост оказывается линейным.

10.3.2 Законы формирования очереди

Рассмотрим процесс обслуживания клиентов некоторой системой, например, выбивание чеков кассой или стрижка в парикмахерской. Иногда для этих процессов употребляют название процесс массового обслуживания. Пусть $\xi(t)$ – число клиентов в зоне обслуживания (считая тех, кто ожидает своей очереди и тех, кто уже обслуживается) в момент времени t . Этот процесс, очевидно, является одним из процессов размножения и гибели. Например, гибелью здесь считается окончание обслуживания и покидание клиентом зоны обслуживания.

Будем считать, что, независимо от числа уже имеющихся клиентов, число прибывающих клиентов имеет одну и ту же интенсивность («клиентам все равно, какой длины очередь»), интенсивность обслуживания клиентов также постоянна («мастер обладает железными нервами и никуда не спешит»). Если канал обслуживания только один, т.е. одновременно не может обслуживаться более одного клиента, то отсюда следует, что для некоторых фиксированных λ, μ в рамках принятой модели ($\forall n$) $\lambda_n = \lambda$, $\mu_n = \mu$.

Обозначим через τ_j время обслуживания j -го клиента. Ранее было показано, что случайные величины τ_j имеют экспоненциальные распределения, более точно

$$(\forall t \geq 0) \mathbf{P}(\tau_j \geq t) = e^{-\mu t}.$$

Если всего имеется N каналов обслуживания (касс, парикмахерских или стоматологических кресел и т.п.), то одно из них освободится через время τ – минимум из величин τ_1, \dots, τ_N . Запишем

$$\mathbf{P}(\tau \geq t) = \prod_{j=1}^N \mathbf{P}(\tau_j \geq t) = e^{-N\mu t}.$$

Мы видим, что время ожидания первого освободившегося канала обслуживания здесь также имеет экспоненциальное распределение. Таким образом, если имеется N каналов обслуживания, то $\mu_j = j\mu$

при $j \leq N$ и $\mu_j = N\mu$ при $j > N$. Это означает, что, после того, как число клиентов достигнет числа каналов обслуживания, что можно интерпретировать, как полный вход системы в рабочий режим, интенсивность гибели соответствующего процесса перестает меняться. Таким образом, систему с произвольным конечным числом каналов в асимптотике можно рассматривать как процесс размножения и гибели с постоянными λ, μ в достаточно естественных предположениях.

Это позволяет вернуться к изучению системы с одним каналом обслуживания. Найдем соответствующее стационарное распределение. В обозначениях, введенных перед (6.5), у нас

$$\pi_j = \left(\frac{\lambda}{\mu}\right)^j, \quad \sum_{j=1}^{\infty} \pi_j = \frac{\mu}{\mu - \lambda}$$

(конечно же, только для $\mu > \lambda$, иначе очередь растет, и ни о каких стационарных распределениях говорить не имеет смысла). Отсюда искомое распределение имеет вид

$$p_n = \frac{\mu - \lambda}{\mu} \left(\frac{\lambda}{\mu}\right)^n, \quad n \geq 0.$$

Поскольку один из клиентов обслуживается, а n в последней формуле равно общему числу имеющихся в зоне обслуживания клиентов, то вероятность того, что длина очереди m равна 0, вычисляется как $p_0 + p_1$, а для $m \geq 1$ совпадает с вероятностью p_{m+1} .

Рассмотрим другой крайний случай – наличие бесконечного числа каналов обслуживания. При этом каждый появляющийся в зоне обслуживания клиент немедленно начинает обслуживаться. Поэтому никакой очереди здесь не образуется. Такую модель системы массового обслуживания иногда называют моделью телефонного узла. Мы уже знаем, что в такой ситуации при произвольном n

$$\lambda_n = \lambda, \quad \mu_n = n\mu,$$

а следовательно,

$$\pi_j = \frac{\lambda^j}{j!\mu^j}, \quad j = 0, 1, \dots; \quad \sum_{j=0}^{\infty} \pi_j = e^{\lambda/\mu},$$

и

$$p_n = \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n e^{-\frac{\lambda}{\mu}}, \quad n \geq 0.$$

10.3. Примеры

Это – вероятность того, что одновременно через телефонный узел ведутся n разговоров. Как мы видим, число одновременно ведущихся разговоров имеет пуассоновское распределение.

Глава 11

О непрерывных модификациях

В этом разделе будем считать, что имеется некоторый конечный сегмент числовой прямой $[a, b]$, на котором заданы все рассматриваемые случайные процессы. Напомним, что реализацией или траекторией случайного процесса называется та неслучайная функция, которая получается при фиксации исхода $\omega \in \Omega$. Это то, что мы получаем, наблюдая развитие изучаемого случайного процесса в конкретном эксперименте. Согласно сложившимся традициям, исходы случайного эксперимента в аргументах случайных величин, как правило, опускаются. Поэтому реализация случайного процесса будет у нас обозначаться $\xi(t)$, т.е. в точности так же, как и сам случайный процесс. Чтобы понять, что именно скрывается за обозначением $\xi(t)$, таким образом, далее придется особенно внимательно следить за контекстом.

Начнем со следующего определения. Два случайных процесса $\xi(t), \eta(t)$ называются *стохастически эквивалентными*, если

$$(\forall t) \mathbf{P}(\xi(t) = \eta(t)) = 1.$$

Любой процесс, стохастически эквивалентный $\xi(t)$, будем называть *модификацией* $\xi(t)$.

Существование модификаций процесса, которые обладают непрерывными реализациями, весьма удобно. Для таких процессов мы имеем возможность исследовать, например, наибольшие и наименьшие значения на отрезках и их распределения как случайных вели-

чин. Дело в том, что, если случайный процесс $\xi(t)$ обладает лишь непрерывными модификациями, то для произвольного x , например имеет место следующее соотношение.

$$\left\{ \sup_{t \in [a, b]} \xi(t) < x \right\} = \bigcap_{t \in \mathbf{Q}[a, b]} \{ \xi(t) < x \},$$

где пересечение берется по множеству всех рациональных чисел отрезка $[a, b]$. Это пересечение счетно, каждое из пересекающихся множеств является случайным событием, а следовательно, и точная верхняя грань значений случайного процесса на отрезке по определению оказывается случайной величиной. При этом возможность обойтись лишь счетным набором множеств в последней формуле напрямую связана с непрерывностью $\xi(t)$. Понятно, что вместо самого случайного процесса в последней формуле можно использовать его модификацию.

Сформулируем две теоремы, в которых формулируются достаточные условия для существования модификации процесса, все реализации которой непрерывны.

Теорема 15. *Пусть найдется такая константа c , что при некоторых $r, \alpha > 0$*

$$(\forall s, t) \mathbf{M}|x(t) - x(s)|^r \leq c|t - s|^{1+\alpha},$$

тогда найдется модификация $\xi(t)$ случайного процесса $x(t)$, все реализации которой непрерывны.

Теорема 16. *Пусть найдены такие монотонно возрастающие при положительных значениях аргументов функции g, y , что*

$$(\forall t, h > 0) \mathbf{P}(|x(t+h) - x(t)| \geq g(h)) \leq y(h),$$

и добавок

$$\sum_{n=0}^{\infty} 2^n y \left(\frac{b-a}{2^n} \right) < \infty, \quad \sum_{n=0}^{\infty} g \left(\frac{b-a}{2^n} \right) < \infty.$$

Тогда найдется модификация $\xi(t)$ случайного процесса $x(t)$, обладающая непрерывными реализациями.

Покажем, что теорема 15 следует из теоремы 16. Действительно, при заданных положительных r, α всегда можно найти такое $\gamma > 0$, что $\alpha - r\gamma > 0$. Положим

$$g(h) = h^\gamma, \quad y(h) = ch^{1+\alpha-r\gamma},$$

где постоянная c также из условия теоремы 15. Из неравенства Чебышева с r -м моментом следует, что

$$\mathbf{P}(|x(t+h) - x(t)| \geq g(h)) \leq \frac{ch^{1+\alpha}}{g^r(h)} = y(h).$$

Осталось проверить сходимость рядов в условии теоремы 16, что сделать совсем несложно.

Для доказательства теоремы 16 нам понадобится следующий вспомогательный результат, оказывающийся полезным и во многих других задачах теории вероятностей, особенно связанных со сходимостью почти наверное.

Лемма 7. (Борель – Кантелли) Пусть для случайных событий A_1, A_2, \dots имеет место сходимость ряда:

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty.$$

Тогда с вероятностью 1 одновременно может произойти лишь конечное число из этих событий.

Доказательство. Используя интерпретацию кванторов существования и всеобщности через теоретико-множественные операции, заметим, что бесконечное число событий данной последовательности происходит лишь тогда, когда происходит

$$A = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n.$$

Введем в рассмотрение монотонно убывающую последовательность событий $B_k = \bigcup_{n \geq k} A_n$. При этом в силу сходимости ряда из вероятностей

$$\mathbf{P}(B_k) \leq \sum_{n=k}^{\infty} \mathbf{P}(A_n) \longrightarrow 0, \quad k \rightarrow \infty.$$

Применяя аксиому непрерывности вероятности, получаем, что

$$\mathbf{P}(A) = \lim_{k \rightarrow \infty} \mathbf{P}(B_k) = 0.$$

Лемма доказана.

Доказательство теоремы 16. Для упрощения формул будем считать, что процессы заданы на $[0,1]$. Рассмотрим случайный процесс

$$X^{(n)}(t) = 2^n (x(t_{n,k})(t_{n,k+1} - t) + x(t_{n,k+1})(t - t_{n,k})),$$

$$t \in [t_{n,k}, t_{n,k+1}],$$

где $t_{n,k} = \frac{k}{2^n}$, $k = 0, \dots, 2^n$. Заметим, что

$$(\forall n, k) X^{(n)}(t_{n,k}) = x(t_{n,k})$$

по построению, а все траектории построенного процесса имеют в качестве графиков непрерывные ломаные. Максимальное значение разности $|X^{(n)}(t) - X^{(n+1)}(t)|$ достигается в одной из 2^n точек вида $\frac{2k-1}{2^{n+1}}$. Привлекая тот факт, что вероятность объединения событий не превосходит суммы их вероятностей, видим, что

$$P_n \equiv \mathbf{P} \left(\max_t |X^{(n)}(t) - X^{(n+1)}(t)| \geq g(2^{-n}) \right) \leq 2^n y(2^{-n}).$$

Из этого неравенства и условий теоремы следует, что числовой ряд $\sum_{n=1}^{\infty} P_n$ мажорируется сходящимся рядом, а значит сходится.

Итак, привлекая лемму Бореля – Кантелли, видим, что произошло лишь конечное число событий

$$A_n = \left\{ \sup_t |X^{(n)}(t) - X^{(n+1)}(t)| \geq g(2^{-n}) \right\},$$

т.е., начиная с некоторого, конечного с вероятностью 1, номера события $N(t)$, они уже не происходят.

С этого момента зафиксируем случайный аргумент ω и будем изучать полученные ломаные и их пределы как неслучайные функции, что позволит задействовать весь аппарат классического математического анализа. В силу сходимости ряда $\sum_{n=1}^{\infty} g(2^{-n})$ мы получаем равномерную относительно t сходимость $\sum_{n=1}^{\infty} (X^{(n)}(t) - X^{(n+1)}(t))$, который, как выяснилось, им мажорируется.

Запишем очевидное соотношение

$$X^{(n)}(t) = X^{(1)}(t) - \sum_{j=1}^{n-1} \left(X^{(j)}(t) - X^{(j+1)}(t) \right).$$

Отсюда следует, что последовательность непрерывных функций $X^{(n)}(t)$ также равномерно сходится относительно t , откуда вытекает, что существует непрерывная функция

$$\hat{X}(t) = \lim_{n \rightarrow \infty} X^{(n)}(t).$$

Вспомним теперь об ω и, вернув его на «законное место» аргумента $\xi(t)$, определим случайный процесс соотношением

$$\xi(t) = \begin{cases} \hat{X}(t), & N(t) < \infty, \\ 0, & N(t) = \infty. \end{cases}$$

Для завершения доказательства осталось проверить, что $\xi(t)$ является модификацией $x(t)$. Второй из случаев в последней формуле реализуется, как мы знаем, только с вероятностью 0, поэтому в дальнейшем доказательстве им можно пренебречь.

При $t = k/2^n$ значения $\xi(t), x(t)$ совпадают с вероятностью 1 по построению. Пусть t не имеет указанного вида. Зафиксируем $\varepsilon > 0$ и рассмотрим

$$w_h = \sup\{|\xi(s) - \xi(q)|, |s - q| < h\}.$$

В силу равномерной непрерывности траекторий процесса $\xi(t)$ справедливо $w_h \rightarrow 0$ при $h \rightarrow 0$ даже с вероятностью 1. Построим последовательность

$$t_m = \frac{k(m)}{2^{n(m)}} \longrightarrow t, \quad m \rightarrow \infty.$$

Тогда

$$\begin{aligned} \mathbf{P}(|x(t) - \xi(t)| \geq \varepsilon) &\leq \mathbf{P}(|x(t) - x(t_m)| \geq \varepsilon/2) + \\ &+ \mathbf{P}(|\xi(t) - \xi(t_m)| \geq \varepsilon/2). \end{aligned}$$

Второе слагаемое в этой сумме не превосходит

$$\mathbf{P}(w_{t-t_m} \geq \varepsilon/2) \rightarrow 0, \quad m \rightarrow \infty.$$

Оценим первое слагаемое. Поскольку $g(t - t_m) \rightarrow 0$, то при достаточно больших m справедливо

$$\begin{aligned} \mathbf{P}(|x(t) - x(t_m)| \geq \varepsilon/2) &\leq \mathbf{P}(|x(t) - x(t_m)| \geq g(t - t_m)) \leq \\ &\leq y(t - t_m) \rightarrow 0, \quad m \rightarrow \infty, \end{aligned}$$

откуда $\mathbf{P}(|x(t) - \xi(t)| \geq \varepsilon) = 0$.

Фактически, доказательство теоремы завершено в силу произвольной малости ε , но, тем не менее, сделаем его более строгим с помощью следующих рассуждений. При каждом натуральном n введем в рассмотрение событие

$$Z_n = \{|x(t) - \xi(t)| \geq 1/n\},$$

которое, как только что было доказано, имеет нулевую вероятность. При этом, очевидно, для произвольного n имеет место включение $Z_n \subset Z_{n+1}$. Осталось заметить, что, по аксиоме непрерывности вероятности,

$$\mathbf{P}(x(t) \neq \xi(t)) = \mathbf{P}\left(\bigcup_{n=1}^{\infty} Z_n\right) = \lim_{n \rightarrow \infty} \mathbf{P}(Z_n) = 0.$$

Теперь теорема доказана полностью.

Глава 12

Условные математические ожидания

В этой главе весьма важным для нас будет понятие измеримости отображения относительно системы множеств. Пусть на множестве Ω выбрана некоторая система подмножеств \mathcal{G} . Напомним, что отображение $f : \Omega \rightarrow \mathbf{R}$ называется \mathcal{G} -измеримым, если прообраз $f^{-1}(B)$ любого борелевского множества B на числовой прямой всегда оказывается элементом \mathcal{G} .

Понятие, вынесенное в заголовок раздела, в современных учебных планах принято относить к теории вероятностей, предмету, который изучается на младших курсах математических факультетов. Но реального применения в этой учебной дисциплине оно чаще всего не находит – слишком много в ней других интересных и важных тем и понятий. Для курса же теории случайных процессов понятия условных математических ожиданий относительно сигма-алгебр или случайных величин крайне важны. Поэтому мы приводим соответствующие определения здесь, считая нелишним напомнить его даже тому читателю, который уже знакомился с ними раньше.

Неформально, любое условное понятие (Z/U) напрямую связано с возможностью получать информацию об одном процессе Z , имея возможность наблюдать развитие некоторого другого процесса U .

Другими словами, то, что находится в условии, мы как бы знаем полностью, имеем полную информацию об U . Тогда условное понятие характеризует ту информацию, которую мы можем получить при этом о процессе Z .

Если выразится немного иначе, условная характеристика Z при условии U это то, что мы можем сказать о Z в терминах U . При этом U предполагается полностью известным.

Начнем с простых вещей. Напомним, что условной вероятностью события A при условии B , если $\mathbf{P}(B) \neq 0$ называется

$$\mathbf{P}(A/B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Если для случайной величины ξ соответствующий интеграл имеет смысл, определим число

$$\mathbf{M}(\xi/B) = \int_{\Omega} \xi d\mathbf{P}(\cdot/B),$$

которое назовем условным математическим ожиданием при условии события B . Тогда

$$\mathbf{M}(\xi/B) = \frac{1}{\mathbf{P}(B)} \int_{\Omega} \xi d\mathbf{P}(\omega \cap B) = \frac{1}{\mathbf{P}(B)} \int_B \xi d\mathbf{P}.$$

Вынося постоянное математическое ожидание за знак интеграла, видим, что

$$\int_B \mathbf{M}(\xi/B) d\mathbf{P} = \mathbf{M}(\xi/B) \cdot \mathbf{P}(B) = \int_B \xi d\mathbf{P}. \quad (12.1)$$

Немного усложним ситуацию. Пусть $\langle \Omega, \mathcal{F}, \mathbf{P} \rangle$ – вероятностное пространство, класс множеств $\mathcal{U} = \{B_1, B_2, \dots\} \subset \mathcal{F}$ осуществляет разбиение Ω , т.е.

$$\bigcup_j B_j = \Omega, \quad i \neq j \Rightarrow B_i \cap B_j = \emptyset,$$

и $(\forall j) \mathbf{P}(B_j) \neq 0$. Случайная величина

$$\hat{\xi} = \sum_j \mathbf{M}(\xi/B_j) \mathbf{1}_{B_j}$$

называется условным математическим ожиданием ξ при условии разбиения \mathcal{U} и обозначается $\mathbf{M}(\xi/\mathcal{U})$. Заметим при этом, что, поскольку $\hat{\xi}$ постоянна на элементах $B_j \in \mathcal{U}$, то она измерима относительно \mathcal{U} .

Пусть $B \in \sigma(\mathcal{U})$. Обозначим C_1, C_2, \dots множества, образующие ту часть разбиения, что

$$B = \bigcup_k B_{i_k} = \bigcup_k C_k.$$

Отметим, что из (12.1) извлекается

$$\int_B \xi d\mathbf{P} = \sum_k \int_{C_k} \xi d\mathbf{P} = \sum_k \int_{C_k} \mathbf{M}(\xi/C_k) d\mathbf{P},$$

иначе говоря, для произвольного $B \in \sigma(\mathcal{U})$ справедливо

$$\int_B \xi d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{U}) d\mathbf{P}$$

(на B все «лишние» индикаторы обращаются в 0).

Для разнообразных доказательств полезным окажется следующее нетрудно проверяемое утверждение. Если имеется такая $\sigma(\mathcal{U})$ -измеримая случайная величина η , что

$$(\forall B \in \sigma(\mathcal{U})) \int_B \eta d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{U}) d\mathbf{P},$$

то $\eta = \mathbf{M}(\xi/\mathcal{U})(\text{mod } \mathbf{P})$, то есть фактически эта величина и есть наше условное математическое ожидание. Дадим теперь самое общее определение.

Пусть \mathcal{G} – σ -алгебра. Тогда \mathcal{G} -измеримая случайная величина $\mathbf{M}(\xi/\mathcal{G})$ называется *условным математическим ожиданием* случайной величины ξ относительно этой σ -алгебры, если для произвольного $B \in \mathcal{G}$ справедливо

$$\int_B \xi d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{G}) d\mathbf{P}.$$

Привлекая сделанное выше замечание, видим, что для поиска условного математического ожидания достаточно указать такую \mathcal{G} -измеримую случайную величину η , что

$$(\forall B \in \mathcal{G}) \int_B \xi d\mathbf{P} = \int_B \eta d\mathbf{P}. \quad (12.2)$$

С учетом того, как вводится скалярное произведение в $\mathbf{L}_2(\Omega)$, равенство из определения условного математического ожидания перепишем в виде

$$(\forall B \in \mathcal{G}) \quad \langle \xi, \mathbf{1}_B \rangle_{\mathbf{L}_2(\Omega)} = \langle \mathbf{M}(\xi/\mathcal{G}), \mathbf{1}_B \rangle_{\mathbf{L}_2(\Omega)} .$$

Если только быть уверенным, что (а это действительно так, см. свойство 2 ниже)

$$\mathbf{M}\xi^2 = \mathbf{M}(\mathbf{M}(\xi/\mathcal{G}))^2 ,$$

то это можно интерпретировать, как равенство углов, образуемых случайной величиной и ее условным математическим ожиданием со всеми порождающими элементами σ -алгебры \mathcal{G} . Если же переписать равенство скалярных произведений в виде

$$\langle \xi - \mathbf{M}(\xi/\mathcal{G}), \mathbf{1}_B \rangle_{\mathbf{L}_2(\Omega)} = 0,$$

то становится понятно, что условное математическое ожидание – это ортогональная проекция ξ на подпространство $\mathbf{L}_2(\Omega)$, натянутое на индикаторы элементов \mathcal{G} .

Существование условного математического ожидания следует из теоремы функционального анализа, формулируемой ниже.

Теорема 17. (Радон - Никодим) Пусть $g(B)$ является σ -аддитивной действительнoзначной функцией на σ -алгебре \mathcal{F} , μ – мера на $\langle \Omega, \mathcal{F} \rangle$. Тогда найдется единственная mod (μ) неотрицательная μ -измеримая функция f такая, что

$$(\forall B \in \mathcal{F}) \quad g(B) = \int_B f(x) d\mu.$$

В этом контексте условное математическое ожидание случайной величины η можно назвать производной σ -аддитивной функции $\int_B \eta d\mathbf{P}$ по мере \mathbf{P} .

Если снова вспомнить неформальное рассуждение, которым открывался настоящий раздел, то условное относительно σ -алгебры \mathcal{G} математическое ожидание случайной величины ξ представляет собой ту составляющую этой случайной величины, которую мы полностью знаем, располагая информацией о каждом из событий, входящих в \mathcal{G} . При этом, поскольку σ -алгебра является, вообще говоря, «достаточно богатым» по содержанию объектом, эта известная часть чаще всего все еще оказывается не постоянной, а представляет собой нетривиальную случайную величину.

Перейдем теперь к формулировкам и доказательствам основных свойств условных математических ожиданий. Пусть, как и выше, \mathcal{G} является σ -алгеброй.

1. Если ξ является \mathcal{G} -измеримой, то $\mathbf{M}(\xi/\mathcal{G}) = \xi$.

Действительно,

$$(\forall B \in \mathcal{G}) \int_B \xi d\mathbf{P} = \int_B \xi d\mathbf{P} \Rightarrow \xi = \mathbf{M}(\xi/\mathcal{G})$$

по определению условного математического ожидания.

2. $\mathbf{M}\mathbf{M}(\xi/\mathcal{G}) = \mathbf{M}\xi$.

Для доказательства этого свойства достаточно выбрать в определении $B = \Omega$:

$$\mathbf{M}\mathbf{M}(\xi/\mathcal{G}) = \int_{\Omega} \mathbf{M}(\xi/\mathcal{G}) d\mathbf{P} = \int_{\Omega} \xi d\mathbf{P} = \mathbf{M}\xi.$$

3. Для произвольных действительных чисел α, β

$$\mathbf{M}(\alpha\xi + \beta\eta/\mathcal{G}) = \alpha\mathbf{M}(\xi/\mathcal{G}) + \beta\mathbf{M}(\eta/\mathcal{G}).$$

Зафиксируем $B \in \mathcal{G}$. Заметим, что правая часть доказываемого равенства \mathcal{G} -измерима как комбинация измеримых величин. Отсюда, привлекая свойство линейности интеграла по вероятности, имеем

$$\begin{aligned} & \int_B (\alpha\mathbf{M}(\xi/\mathcal{G}) + \beta\mathbf{M}(\eta/\mathcal{G})) d\mathbf{P} = \\ & = \alpha \int_B \xi d\mathbf{P} + \beta \int_B \eta d\mathbf{P} = \int_B (\alpha\xi + \beta\eta) d\mathbf{P}, \end{aligned}$$

что и требовалось доказать.

4. Если ξ не зависит от \mathcal{G} , то $\mathbf{M}(\xi/\mathcal{G}) = \mathbf{M}\xi$.

Смысл этого свойства очевиден – если мы располагаем информацией о том, что не связано с ξ , то ее часть, которую мы можем описать, тривиальна. Докажем свойство.

Понятно, что

$$(\forall B \in \mathcal{G})(\forall x) \quad \mathbf{P}(\xi < x, B) = \mathbf{P}(\xi < x)\mathbf{P}(B),$$

т.е. ξ и $\mathbf{1}_B$ независимы. Отсюда

$$\int_B \xi d\mathbf{P} = \mathbf{M}\xi \mathbf{1}_B = \mathbf{M}\xi \mathbf{M}\mathbf{1}_B = \mathbf{M}\xi \int_B d\mathbf{P} = \int_B \mathbf{M}\xi d\mathbf{P},$$

что завершает доказательство.

5. Если η \mathcal{G} -измерима, то для произвольной ξ

$$\mathbf{M}(\xi\eta/\mathcal{G}) = \eta\mathbf{M}(\xi/\mathcal{G}).$$

Таким образом, при изучении условных математических ожиданий с величинами, измеримыми относительно \mathcal{G} можно работать как с константами. Докажем это.

Пусть $\eta = \mathbf{1}_C$, $C \in \mathcal{G}$. Тогда для $B \in \mathcal{G}$

$$\begin{aligned} \int_B \eta\mathbf{M}(\xi/\mathcal{G})d\mathbf{P} &= \int_{B \cap C} \mathbf{M}(\xi/\mathcal{G})d\mathbf{P} = \\ &= \int_{B \cap C} \xi d\mathbf{P} = \int_B \xi\eta d\mathbf{P}, \end{aligned}$$

т.е. в этой ситуации все доказано. Случай линейных комбинаций индикаторов очевиден, общий случай получается предельным переходом.

6. Пусть $\mathcal{G} \subset \mathcal{F}$ – сигма-алгебры. Тогда

$$\mathbf{M}(\mathbf{M}(\xi/\mathcal{F})/\mathcal{G}) = \mathbf{M}(\xi/\mathcal{G}).$$

Действительно, при $B \in \mathcal{G}$ справедливо

$$\begin{aligned} \int_B \mathbf{M}(\mathbf{M}(\xi/\mathcal{F})/\mathcal{G}) d\mathbf{P} &= \int_B \mathbf{M}(\xi/\mathcal{F})d\mathbf{P} = \\ &= \int_B \xi d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{G})d\mathbf{P}, \end{aligned}$$

что и заканчивает доказательство.

7. Если $\mathcal{G} \subset \mathcal{F}$, η – \mathcal{F} -измерима, то

$$\mathbf{M}(\xi\eta/\mathcal{G}) = \mathbf{M}(\eta\mathbf{M}(\xi/\mathcal{F})/\mathcal{G}).$$

Доказательство следует из свойств 6 и 5.

Глава 13

Мартингалы

13.1 Определения

В заключительных разделах курса изучим случайные функции, имеющие зависимые приращения и устроенные иначе, чем марковские процессы. Будем считать, что все рассматриваемые ниже процессы заданы на ограниченном полуинтервале $T = [a, b)$, его замыкании или каком-либо дискретном подмножестве этого полуинтервала.

Будем говорить, что семейство σ -алгебр $\mathcal{F}_t, t \in T$ образует *поток*, если при $s < t$ справедливо $\mathcal{F}_s \subset \mathcal{F}_t$. Поток будем называть непрерывным справа в точке a , если σ -алгебры \mathcal{F}_a и $\mathcal{F}_{a+} = \bigcap \{\mathcal{F}_t, t > a\}$ совпадают. Заметим, что σ -алгебры $\mathcal{F}_{t+}, t \in T$ образуют поток, непрерывный справа в любой точке T .

Если при каждом t случайная величина $\xi(t)$ \mathcal{F}_t -измерима, то будем говорить, что случайный процесс $\xi(t)$ *согласован с потоком*. Очевидно, что любой случайный процесс $\xi(t)$ согласован с потоком $\mathcal{F}_t^{\xi} = \sigma(\xi(s), s \leq t)$.

Случайный процесс называется *мартингалом* относительно потока сигма-алгебр, если он согласован с этим потоком, имеет конечное математическое ожидание, и

$$(\forall t, s \in T) (s < t) \Rightarrow \mathbf{M}(\xi(t)/\mathcal{F}_s) = \xi(s)$$

с вероятностью 1. Если $\mathbf{M}(\xi(t)/\mathcal{F}_s) \geq \xi(s)$ для произвольных пар аргументов с условием $s < t$, то процесс называется *субмартингалом*, если же в последнем определении знак \geq изменен на \leq , то

13.1. Определения

процесс $\xi(t)$ называют *супермартингалом*. Обобщающим понятием служит *полумартингал*, так называют субмартингал или супермартингал.

Как мы видим, при упоминании мартингала формально требуется описать поток σ -алгебр, относительно которого рассматриваемый процесс является мартингалом. Тем не менее, часто этого не делается. Обычно считают, что, если поток σ -алгебр отдельно не упомянут, то имеется в виду «естественный» поток $\mathcal{F}_t^\xi = \sigma(\xi(s), s \leq t)$.

Приведем здесь три примера мартингалов.

1. Пусть $\xi(t)$ – процесс с независимыми приращениями и постоянным математическим ожиданием. Тогда это мартингал относительно \mathcal{F}_t^ξ , $t \in T$. Действительно, при $s < t$

$$\mathbf{M}(\xi(t)/\mathcal{F}_s^\xi) = \mathbf{M}((\xi(t) - \xi(s))/\mathcal{F}_s^\xi) + \mathbf{M}(\xi(s)/\mathcal{F}_t^\xi),$$

но, согласно свойствам условных математических ожиданий, первое слагаемое здесь равно $\mathbf{M}(\xi(t) - \xi(s)) = 0$, а второе совпадает с $\xi(s)$.

2. Пусть ξ – случайная величина и задан поток σ -алгебр \mathcal{F}_t , $t \in T$. Тогда $X(t) = \mathbf{M}(\xi/\mathcal{F}_t)$ – мартингал. Это вытекает из следующей цепочки равенств ($s < t$):

$$\mathbf{M}(X(t)/\mathcal{F}_s) = \mathbf{M}(\mathbf{M}(\xi/\mathcal{F}_t)/\mathcal{F}_s) = \mathbf{M}(\xi/\mathcal{F}_s) = X(s).$$

Конечно же, снова были использованы свойства условных математических ожиданий из предыдущего раздела. Заметим также, что здесь рассмотренный процесс является мартингалом относительно потока \mathcal{F}_t , $t \in T$. Совпадения этих σ -алгебр с элементами «естественного» потока может и не наблюдаться.

3. Последовательность сумм независимых одинаково распределенных случайных величин также образует мартингал (с дискретным временем). Это немедленно получается из рассуждений первого примера и уже отмечавшегося ранее того факта, что такие суммы образуют процесс с дискретным временем и независимыми приращениями.

Очевидно, что определение субмартингала можно переписать в следующем виде

$$(\forall s, t)(\forall B \in \mathcal{F}_s) (s < t) \Rightarrow \int_B \xi(s) d\mathbf{P} \leq \int_B \xi(t) d\mathbf{P}. \quad (13.1)$$

Теорема 18. Пусть $\xi(t)$ – мартингал, $f(\cdot)$ – измеримая функция, выпуклая вниз и такая, что $\mathbf{M}f(\xi(t)) < \infty$. Тогда $f(\xi(t))$ субмартингал относительно того же потока σ -алгебр.

Доказательство. Согласованность с первоначальным потоком очевидна. Проверим условие (13.1) для $f(\xi(t))$. Из выпуклости вниз следует существование в каждой точке графика функции опорной прямой, т.е.

$$(\forall x_0)(\exists C)(\forall x) \quad f(x) \geq f(x_0) + C(x_0) \cdot (x - x_0).$$

Фиксируем $N \in \mathbf{N}$, $B \in \mathcal{F}_s$ пусть $B_N = B \cap \{|C(\xi(s))| \leq N\}$. Тогда

$$\int_{B_N} f(\xi(t)) d\mathbf{P} \geq \int_{B_N} f(\xi(s)) d\mathbf{P} + \int_{B_N} C(\xi(s))(\xi(t) - \xi(s)) d\mathbf{P}. \quad (13.2)$$

Заметим, что при $t < s$ обязательно $B_N \in \mathcal{F}_s$ и

$$\mathbf{M}(C(\xi(s)) \cdot (\xi(t) - \xi(s)) / \mathcal{F}_s) = C(\xi(s)) \mathbf{M}((\xi(t) - \xi(s)) / \mathcal{F}_s) = 0,$$

откуда следует, что второй интеграл в (13.2) равен 0. Таким образом,

$$\int_{B_N} f(\xi(t)) d\mathbf{P} \geq \int_{B_N} f(\xi(s)) d\mathbf{P}.$$

Доказательство теоремы немедленно получается из этого неравенства предельным переходом при $N \rightarrow \infty$.

Простым следствием теоремы является возможность строить многочисленные примеры субмартингалов. Например, если $\xi(t)$ – мартингал, то $|\xi(t)|$, $\xi^2(t)$, $e^{\xi(t)}$ – субмартингалы. Сколько угодно примеров супермартингалов можно построить, если иметь ввиду несложно проверяемый вариант доказанной теоремы с функцией, выпуклой вверх.

Оказывается, между мартингалами и субмартингалами имеется следующая интересная связь.

13.1. Определения

Теорема 19. Пусть последовательность X_n , $n \in \{0\} \cup \mathbf{N}$ является субмартингалом относительно потока σ -алгебр \mathcal{F}_n . Тогда найдутся мартингал ξ_n и возрастающая случайная последовательность A_n , согласованная с тем же потоком такая, что $(\forall n)$ A_n измерима относительно \mathcal{F}_{n-1} и

$$(\forall n \in \{0\} \cup \mathbf{N}) \quad X_n = \xi_n + A_n \quad (13.3)$$

с вероятностью единица.

Доказательство. Пусть $\xi_0 = X_0$, $A_0 = 0$, а при каждом натуральном n определим

$$\begin{aligned} \xi_n &= \xi_0 + \sum_{j=0}^{n-1} (X_{j+1} - \mathbf{M}(X_{j+1}/\mathcal{F}_j)), \\ A_n &= \sum_{j=0}^{n-1} (\mathbf{M}(X_{j+1}/\mathcal{F}_j) - X_j). \end{aligned}$$

Требуемые свойства измеримости здесь легко проверяются. Теорема доказана.

Представление субмартингала в виде (13.3) иногда называют *разложением Дуба*. В [7, с. 475 – 476] доказана также единственность подобного разложения. Последовательность A_n называют *компенсирующей последовательностью* для субмартингала X_n . В силу ее измеримости относительно «предыдущей» σ -алгебры ее принято называть предсказуемой.

Много дополнительных фактов о полумартингалах можно найти в книгах [12], [14] и [7]. Нам потребуются следующие два неравенства из [7, с. 485] (частный случай теоремы Дуба для субмартингалов). Их доказательства опустим. Пусть

$$\eta_u = \sup_{a \leq t \leq u} \xi(t).$$

Теорема 20. Неотрицательный субмартингал $\xi(t)$, все траектории которого непрерывны справа, для каждого $u \in (a, b)$ удовлетворяет неравенствам

$$\begin{aligned} \mathbf{P}(\eta_u \geq x) &\leq \frac{\mathbf{M}\xi(u)}{x}, \\ \mathbf{M}\eta_u^p &\leq \left(\frac{p}{p-1}\right)^p \mathbf{M}\xi^p(u), \quad p > 1. \end{aligned}$$

13.2 Полнота пространства мартингалов

Пусть $\langle \Omega, \mathcal{F}, \mathbf{P} \rangle$ – вероятностное пространство, поток σ -алгебр \mathcal{F}_t , $t \in T = [a, b]$ непрерывен в каждой точке справа. Предположим также, что все σ -алгебры, входящие в поток, полны относительно \mathbf{P} , т.е. содержат все возможные подмножества своих событий нулевой вероятности. Введем в рассмотрение класс $\mathcal{M}(T)$ всех мартингалов относительно этого потока, у которых реализации с вероятностью 1 не имеют разрывов второго рода и

$$\mathbf{M}\xi^2(t) < \infty, t \in T.$$

Пусть $\mathcal{M}_C(T)$ – подкласс мартингалов из $\mathcal{M}(T)$, обладающих непрерывными реализациями с вероятностью 1.

Все значения мартингала $\xi(t)$, $t \in [a, b]$ восстанавливаются по $\xi(b)$. Поэтому нет ничего удивительного, что некоторые понятия определяются здесь только через значение мартингала в этой граничной точке. Так, введем в $\mathcal{M}(T)$ скалярное произведение по формуле

$$\langle \xi, \eta \rangle_T = \mathbf{M}\xi(b)\eta(b).$$

Теорема 21. *$\mathcal{M}(T)$ является гильбертовым пространством, а его подпространство $\mathcal{M}_C(T)$ является замкнутым в норме, индуцированной введенным скалярным произведением.*

Доказательство. Пусть $\xi_n(t)$ – фундаментальная последовательность в $\mathcal{M}(T)$. Тогда последовательность $\xi_n(b)$ в пространстве $\mathbf{L}_2(\langle \Omega, \mathcal{F}, \mathbf{P} \rangle)$ сходится к некоторой случайной величине μ . Положим

$$\mu(t) = \mathbf{M}(\mu/\mathcal{F}_t).$$

Этот случайный процесс является мартингалом, ограничения, наложенные на поток \mathcal{F}_t , $t \in T$, позволяют выбрать модификацию этого случайного процесса, не имеющую разрывов второго рода. Это доказывает первое утверждение теоремы.

Как мы знаем, $|\xi_n(t) - \mu(t)|$ – субмартингал. Из второго неравенства теоремы 20 следует

$$\mathbf{M} \left(\sup_t |\xi_n(t) - \mu(t)| \right)^2 \leq 4\mathbf{M}(\xi_n(b) - \mu(b))^2 \rightarrow 0.$$

13.3. Стохастический интеграл по мартингалу

Таким образом, для каждого натурального k может быть указан номер $n(k)$, для которого

$$\mathbf{M} \sup |\xi_{n(k)}(t) - \mu(t)| \leq \sqrt{\mathbf{M} \sup |\xi_{n(k)}(t) - \mu(t)|^2} < \frac{1}{k^2},$$

а значит, используя первое неравенство теоремы 20, для произвольного $\varepsilon > 0$ выводим

$$\begin{aligned} & \sum_{k=1}^{\infty} \mathbf{P} (\sup |\xi_{n(k)}(t) - \mu(t)| \geq \varepsilon) \leq \\ & \leq \sum_{k=1}^{\infty} \frac{\mathbf{M} \sup |\xi_{n(k)}(t) - \mu(t)|}{\varepsilon} < \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty. \end{aligned}$$

Из критерия сходимости почти наверное, таким образом, следует, что для каждого t реализации $\xi_{n(k)}(t)$ сходятся к реализации $\mu(t)$ почти наверное. В силу ограниченности основного параметрического отрезка $[a, b]$ имеет место равномерная относительно t сходимость. Следовательно, если первоначально выбирались мартингалы с непрерывными реализациями (элементы $\mathcal{M}_C(T)$), то и почти все реализации предельного мартингала непрерывны. Если заменить все разрывные реализации предельного мартингала непрерывными функциями (а это не нарушает его согласованности с потоком, так как все σ -алгебры потока полны), то получается непрерывная модификация процесса $\mu(t)$. Следовательно, пределом последовательности элементов $\mathcal{M}_C(T)$ является мартингал этого же пространства. Теорема доказана.

13.3 Стохастический интеграл по мартингалу

Пусть у нас имеется действительногозначный случайный процесс $\xi(t)$, заданный на $[a, b]$. Назовем его *прогрессивно измеримым* относительно потока σ -алгебр \mathcal{F}_t , $t \in [a, b]$, если при любом t его сужение на множество $[a, t]$ измеримо относительно $\sigma(\mathcal{F}_t \times \mathcal{B}[a, t])$.

Прогрессивная измеримость требует большего, чем просто согласованность случайного процесса с потоком сигма-алгебр. Например, если интерпретировать каждую из \mathcal{F}_t как совокупность

информации, которой располагает наблюдатель к моменту времени t , то согласованность с потоком означает полную информированность наблюдателя к моменту времени t о случайной величине $\xi(t)$. В то же время прогрессивная измеримость – это возможность получать полную информацию о событиях, связанные с поведением процесса не только в одном или счетном числе моментов времени, но и на целых интервалах времени. Например, такая возможность будет означать разрешение оперировать с интегралами от траекторий случайных процессов, которые хотя бы формально становятся нужным образом измеримыми.

Итак, прогрессивно измеримый процесс в частности обладает тем свойством, что для любого t случайная величина $\xi(t)$ измерима относительно \mathcal{F}_t . Оказывается, что при минимальных дополнительных ограничениях ничего дополнительного для появления у процесса свойства прогрессивной измеримости не требуется. Вот один из вариантов подобных не очень ограничительных добавочных требований.

Теорема 22. Пусть процесс $\xi(t)$ согласован с \mathcal{F}_t . Если этот процесс имеет непрерывные справа (или слева) реализации, то он прогрессивно измерим относительно потока σ -алгебр \mathcal{F}_t , $t \in [a, b]$.

Доказательство. Докажем теорему для случая, когда реализации непрерывны слева. Выберем последовательность разбиений нашего отрезка

$$a = t_0 < t_1(n) < \dots < t_n(n) = b$$

так, что максимальная длина отрезка разбиения стремится к нулю при $n \rightarrow \infty$. Пусть

$$\xi_n(t) = \sum_{j=1}^n \xi(t_j(n)) \mathbf{1}_{\{t_{j-1}(n) \leq t < t_j(n)\}}.$$

Тогда при $n \rightarrow \infty$ и любых t, ω выполнено $\xi_n(t) \rightarrow \xi(t)$. Поскольку все функции $\xi_n(t, \omega)$ измеримы относительно $\sigma(\mathcal{F}_t \times \mathcal{B}[a, t])$, то это же относится и к их пределу. Теорема доказана.

Пусть $\mu(t)$, $t \in [0, b]$ – квадратично интегрируемый мартингал. Согласно теореме 18, его квадрат является субмартингалом. Построим для этого субмартингала разложение Дуба, действуя аналогично утверждению теоремы 19. Тогда можно утверждать, что существует неубывающий случайный процесс $A(t)$ такой, что процесс $\mu^2(t) - A(t)$ является мартингалом. Условимся говорить, что

13.3. Стохастический интеграл по мартингалу

$\mu(t)$ имеет абсолютно непрерывную характеристику, если существует прогрессивно измеримый неотрицательный процесс $a(t)$, такой, что

$$(\forall t) \quad A(t) = \int_0^t a(s) ds.$$

При этом под характеристикой мартингала как раз и будем понимать процесс $A(t)$, который для мартингаловых последовательностей выше назывался компенсирующим.

Будем рассматривать квадратично интегрируемый мартингал $\mu(t)$ с абсолютно непрерывной характеристикой относительно непрерывного справа потока σ -алгебр, составляющие которого полны. Дополнительно потребуем, чтобы почти все реализации этого мартингала были бы непрерывны. Легко видеть, что при $t > s$

$$\begin{aligned} \mathbf{M}((\mu(t) - \mu(s))^2 / \mathcal{F}_s) &= \mathbf{M}(\mu^2(t) / \mathcal{F}_s) - \mu^2(s) = \\ &= \mathbf{M}((A(t) - A(s)) / \mathcal{F}_s). \end{aligned}$$

Оказывается, полученное соотношение является необходимым и достаточным для того, чтобы мартингал $\mu(t)$ имел бы случайный процесс $A(t)$ в качестве своей характеристики.

Лемма 8. Пусть $\xi(t)$ — мартингал относительно потока сигма-алгебр \mathcal{F}_t , $t \in [a, b]$, $A(t)$ — неубывающий случайный процесс, измеримый относительно того же потока. Тогда $A(t)$ является характеристикой $\xi(t)$ тогда и только тогда, когда при $t > s$

$$\mathbf{M}(\xi^2(t) / \mathcal{F}_s) - \xi^2(s) = \mathbf{M}(A(t) - A(s) / \mathcal{F}_s). \quad (13.4)$$

Доказательство. Необходимость, по сути очевидную, мы проверили непосредственно перед формулировкой леммы. Проверим достаточность. Из (13.4) немедленно следует, что

$$\mathbf{M}(\xi^2(t) - A(t) / \mathcal{F}_s) = \xi^2(s) + \mathbf{M}(A(t) - A(s) / \mathcal{F}_s) - \mathbf{M}(A(t) / \mathcal{F}_s).$$

Привлекая тот факт, что условное математическое ожидание $A(s)$ относительно \mathcal{F}_s совпадает с ним самим в силу имеющихся свойств измеримости, видим, что

$$\mathbf{M}(\xi^2(t) - A(t) / \mathcal{F}_s) = \xi^2(s) - A(s).$$

Это завершает доказательство леммы.

Приступим к определению интеграла по мартингалу. Сначала определим его для ступенчатых случайных процессов. Разобьем наш отрезок неслучайными точками

$$a = t_0 < t_1 < \dots < t_N = b$$

и рассмотрим случайный процесс

$$\eta(t) = \sum_{j=1}^N \eta_j \mathbf{1}_{\{t_{j-1} \leq t < t_j\}},$$

где случайные величины η_j измеримы относительно \mathcal{F}_{t_j} , $j = 1, \dots, N$ и ограничены с вероятностью 1.

Приступая к определению нового интеграла, введем для него альтернативное обозначение

$$\int_a^t \eta(s) d\mu(s) \equiv \eta \circ \mu(t).$$

Это новое обозначение подчеркивает, что, в силу наличия прогрессивной измеримости у $\eta(t)$, определяемый интеграл также будет случайным процессом, возникающим в результате некоторой операции над процессами $\eta(t), \mu(t)$.

Положим для простого случайного $\eta(t)$ по определению

$$\eta \circ \mu(t) = \sum_{k \leq m(t)} \eta_k (\mu(t_{k+1}) - \mu(t_k)) + \eta_{m(t)} (\mu(t) - \mu(t_{m(t)})),$$

где $m(t) = \max\{k \mid t_k \leq t\}$. Ясно, что для ступенчатого процесса $\eta(t)$ введенная конструкция имеет тип интегральной суммы. Именно этим и объясняется предложенное обозначение ее в виде интеграла.

Случайный процесс $\eta \circ \mu(t)$ имеет непрерывные реализации и является квадратично интегрируемым мартингалом относительно того же потока σ -алгебр, что и μ . Действительно, если k, s таковы, что $t_k \geq s$, то

$$\begin{aligned} & \mathbf{M}(\eta_k(\mu(t_{k+1}) - \mu(t_k)) / \mathcal{F}_s) = \\ & = \mathbf{M}(\mathbf{M}(\eta_k(\mu(t_{k+1}) - \mu(t_k)) / \mathcal{F}_{t_k}) / \mathcal{F}_s) = \\ & = \mathbf{M}(\eta_k \mathbf{M}((\mu(t_{k+1}) - \mu(t_k)) / \mathcal{F}_{t_k}) / \mathcal{F}_s) = 0. \end{aligned}$$

13.3. Стохастический интеграл по мартингалу

Аналогичным образом можно проверить, что при $t_k < s < t$

$$\mathbf{M}(\eta_k(\mu(t) - \mu(t_k))/\mathcal{F}_s) = \eta_k(\mu(s) - \mu(t_k)).$$

Поэтому из предложенного определения нового интеграла для элементарного процесса $\eta(t)$ вытекает равенство

$$\mathbf{M}(\eta \circ \mu(t)/\mathcal{F}_s) = \eta \circ \mu(s), \quad s < t.$$

Пусть $s < t < u < v$, и имеются две произвольных случайных величины, таких, что η измерима относительно \mathcal{F}_s , а $\hat{\eta}$ — относительно \mathcal{F}_u . Тогда

$$\begin{aligned} & \mathbf{M}(\eta(\mu(t) - \mu(s))\hat{\eta}(\mu(v) - \mu(u))/\mathcal{F}_s) = \\ & = \mathbf{M}(\mathbf{M}(\eta(\mu(t) - \mu(s))\hat{\eta}(\mu(v) - \mu(u))/\mathcal{F}_u)/\mathcal{F}_s) = \\ & \mathbf{M}(\eta(\mu(t) - \mu(s))\hat{\eta}\mathbf{M}((\mu(v) - \mu(u))/\mathcal{F}_u)/\mathcal{F}_s) = 0. \end{aligned}$$

Из этой формулы следует, что

$$\mathbf{M}((\eta \circ \mu(t))^2/\mathcal{F}_s) - (\eta \circ \mu(s))^2 = \mathbf{M}\left(\int_s^t \eta^2(u)a(u) du/\mathcal{F}_s\right),$$

т.е., согласно лемме 8, мартингал $\eta \circ \mu$ также обладает абсолютно непрерывной характеристикой.

Если реализации процесса $\eta(t)$ непрерывны слева и ограничены неслучайной постоянной, то интеграл от него можно определить с помощью предельного перехода. Точнее, пусть

$$\eta_n(t) = \sum_{k=0}^{n-1} \eta\left(a + (b-a)\frac{k}{n}\right) \mathbf{1}_{\{a+(b-a)\frac{k}{n} \leq t < a+(b-a)\frac{k+1}{n}\}},$$

тогда, по определению,

$$\eta \circ \mu(t) = \lim_{n \rightarrow \infty} \eta_n \circ \mu(t).$$

Теорема 16 позволяет понимать этот предельный переход как равномерную по t сходимости почти всех реализаций стохастического интеграла $\eta_n \circ \mu(t)$. Действительно,

$$\begin{aligned} \mathbf{M}(\eta_n \circ \mu(b) - \eta_m \circ \mu(b))^2 &= \mathbf{M}((\eta_n - \eta_m) \circ \mu(b))^2 = \\ &= \mathbf{M} \int_a^b (\eta_n(m) - \eta_m(u))^2 a(u) du \longrightarrow 0. \end{aligned}$$

Аналогичным образом, с помощью предельного перехода, определяется стохастический интеграл по мартингалу для всех процессов $\eta(t)$, которые являются пределами последовательностей непрерывных слева ограниченных прогрессивно измеримых процессов в том смысле, что

$$\mathbf{M} \int_a^b (\eta(t) - \eta_n(u))^2 a(u) du \longrightarrow 0, \quad n \rightarrow \infty.$$

Класс таких процессов совпадает с классом прогрессивно измеримых процессов, для которых

$$\mathbf{M} \int_a^b \eta^2(t) a(t) dt < \infty.$$

Единственным препятствием непосредственному определению здесь служит возможность таким процессам иметь неограниченные траектории. Чтобы устранить это несложное затруднение, рассмотрим «урезанный» на уровне u случайный процесс, для которого формально все в порядке:

$$\eta^u(t) = \eta(t) \cdot \mathbf{1}_{\{|\eta(t)| \leq u\}}.$$

Построив для $\eta(t)$ аппроксимирующую последовательность ступенчатых процессов $\eta_n^u(t)$, $n \in \mathbf{N}$, которая, очевидно, без ограничения общности, может считаться неизменной по отношению к различным $u > 0$, полагают

$$\eta \circ \mu(t) = \lim_{n \rightarrow \infty} \lim_{u \rightarrow \infty} \eta_{n,u} \circ \mu(t),$$

где

$$\eta_{n,u}(t) = \int_a^b \eta_n^u(s) \exp\{-(s-t)n\} ds.$$

Конечно же, в процессе определения этого интеграла необходимо убедиться, что результат перехода к пределу не зависит от выбора аппроксимирующей последовательности ступенчатых процессов. Это делается стандартным для теории интегрирования образом.

Глава 14

Стохастические дифференциалы и интегралы Ито

14.1 Стохастические интегралы Ито

Пусть выполнены все предположения предыдущего раздела относительно потока σ -алгебр и прогрессивной измеримости процесса $\xi(t)$. Но теперь вместо произвольного мартингала $\mu(t)$ рассмотрим винеровский процесс $w(t)$. Для него соответствующая элементарная стохастическая мера

$$\mu([a, b]) = w(b) - w(a)$$

обладает тем свойством, что для непересекающихся отрезков Δ_1 и Δ_2 случайные величины $\mu(\Delta_1)$ и $\mu(\Delta_2)$ независимы, а $\xi(a)$ не зависит от $\mu([a, b])$ в силу сделанных предположений. Интеграл

$$\int_a^b \xi(t) dw(t)$$

называют *стохастическим интегралом Ито*. Для непрерывного в среднем квадратическом процессе $\xi(t)$, очевидно, справедливо

$$\int_a^b \xi(s)dw(s) = \lim_{n \rightarrow \infty} \text{l.i.m.} \sum_{k=1}^n \xi(t_{k-1})(w(t_k) - w(t_{k-1})), \quad (14.1)$$

где t_k , $k = 0, \dots, n$ – произвольное разбиение отрезка $[a, b]$, выбранное так, что

$$\Delta = \max_k (t_k - t_{k-1}) \rightarrow 0, \quad n \rightarrow \infty.$$

Нетрудно также проверить, что если $\xi(t)$ не обращается в 0 с вероятностью 1, то для достаточно малых h в среднеквадратическом смысле

$$\int_t^{t+h} \xi(s)dw(s) = \xi(t)(w(t+h) - w(t)) + o(\sqrt{h}).$$

Вычислим, например, интеграл Ито непосредственно по определению (14.1), если в качестве $\xi(t)$ выбран винеровский процесс. Используем равенство

$$2w(t_{k-1})(w(t_k) - w(t_{k-1})) = w^2(t_k) - w^2(t_{k-1}) - (w(t_k) - w(t_{k-1}))^2.$$

Тогда выражение под знаком предела в (14.1) превращается в

$$S_n = \frac{1}{2} (w^2(b) - w^2(a)) - \frac{1}{2} \sum_{k=1}^n (w(t_k) - w(t_{k-1}))^2. \quad (14.2)$$

Докажем, что

$$\lim_{n \rightarrow \infty} \text{l.i.m.} \sum_{k=1}^n (w(t_k) - w(t_{k-1}))^2 = b - a. \quad (14.3)$$

Для этого обозначим

$$\delta_k = (w(t_k) - w(t_{k-1}))^2 - (t_k - t_{k-1}), \quad k = 1, \dots, n.$$

Тогда

$$\sum_{k=1}^n (w(t_k) - w(t_{k-1}))^2 - (b - a) = \sum_{k=1}^n \delta_k,$$

14.2. Стохастический дифференциал

причем случайные величины δ_k независимы, $\mathbf{M}\delta_k = 0$, и

$$\mathbf{M}\delta_k^2 = \mathbf{M}(w(t_k) - w(t_{k-1}))^2 = (t_k - t_{k-1})^2 = 2(t_k - t_{k-1})^2,$$

что следует из того, что для случайной величины ξ , имеющей нормальное распределение с параметрами 0 и σ^2 справедливо $\mathbf{M}\xi^4 = 3\sigma^4$. Отсюда

$$\mathbf{M} \left(\sum_{k=1}^n \delta_k \right)^2 = 2 \sum_{k=1}^n (t_k - t_{k-1})^2 \leq 2\Delta \sum_{k=1}^n (t_k - t_{k-1}) = 2(b-a)\Delta,$$

что стремится к 0 вместе с Δ , т.е. (14.3) доказано. Окончательно, из (14.2) и (14.3),

$$\int_a^b w(s)dw(s) = \frac{1}{2} (w^2(b) - w^2(a) - (b-a)).$$

Особенно привлекательно эта формула выглядит для случая интеграла по $[0, t]$:

$$\int_0^t w(s)dw(s) = \frac{1}{2} (w^2(t) - t). \quad (14.4)$$

Несколько неожиданное на первый взгляд появление t в ответе объясняется тем, что в результате интегрирования мартингала должно получиться мартингал. Но квадрат винеровского процесса мартингалом не является, и появившееся t здесь играет роль компенсирующей функции из разложения Дуба. Проверьте это в качестве простого упражнения.

14.2 Стохастический дифференциал

Напомним, что у нас рассматриваются только случайные процессы, значения которых при всех допустимых t являются элементами $\mathbf{L}_2(\Omega)$. Говорят, что случайный процесс $\xi(t)$ имеет *стохастический дифференциал*

$$d\xi(t) = a(t)dt + b(t)dw(t), \quad (14.5)$$

где $a(t) = a(t, \xi(t))$, $b(t) = b(t, \xi(t))$ – прогрессивно измеримые случайные процессы, если

$$\xi(t) = \xi(t_0) + \int_{t_0}^t a(s) ds + \int_{t_0}^t b(s) dw(s), \quad t \geq t_0, \quad (14.6)$$

а выписанные интегралы существуют.

Поскольку

$$\int_s^{s+h} \sqrt{\mathbf{M}a^2(t)} dt \rightarrow 0, \quad \int_s^{s+h} \sqrt{\mathbf{M}b^2(t)} dt \rightarrow 0$$

при $h \rightarrow 0$ и произвольном s , то нетрудно проверить, что процесс, имеющий стохастический дифференциал, непрерывен в среднем квадратическом.

Оказывается, процессы $a(t, \xi(t))$ и $b(t, \xi(t))$, фигурирующие в (14.5), могут быть интерпретированы следующим образом. Если представить себе, что $\xi(t)$, $t \geq t_0$ описывает движение некоторой частицы, то при условии, что $\xi(s) = x$, за малое время h математическое ожидание смещения этой частицы будет равно $a(s, x)h + o(h)$, а само это смещение с точностью до $o(\sqrt{h})$ будет случайной величиной, имеющей нормальное распределение с дисперсией $b(s, x)$. Поэтому $a(t)$ иногда называют *коэффициентом сноса*, а $b(t)$ – *коэффициентом диффузии*. Процесс, имеющий стохастический дифференциал вида (14.5), называют *диффузионным*. Уравнение (14.5) относительно $\xi(t)$ называют *стохастическим дифференциальным уравнением*. Простое изложение основных моментов теории стохастических дифференциальных уравнений можно найти в [20].

В качестве примера проверим следующее соотношение

$$dw^2(t) = dt + 2w(t)dw(t).$$

Действительно, из (14.4) получаем

$$w^2(t) = w^2(0) + t + 2 \int_0^t w(s)dw(s),$$

откуда, привлекая (14.6), и получается требуемое соотношение. Таким образом, квадрат винеровского процесса является диффузионным процессом с коэффициентом сноса 1 и коэффициентом диффузии $2w(t)$.

14.2. Стохастический дифференциал

В физике стохастические дифференциальные уравнения чаще всего изучают в рамках волновой теории. Основным методом их решения является поиск решения в виде плотности некоторого вероятностного распределения и преобразованием первоначального уравнения в так называемое уравнение Фоккера-Планка. Уравнение Фоккера-Планка – неслучайное дифференциальное уравнение в частных производных. Оно определяет временную эволюцию плотности распределения так же, как уравнение Шрёдингера определяет зависимость волновой функции системы от времени в квантовой механике или уравнение диффузии в химии задает изменение концентрации какого-либо реагента со временем. Кроме этого, иногда решения можно искать численно, например с помощью метода Монте-Карло. Другие техники нахождения решений используют специальные приемы вычисления интегралов по контурам (подробнее см. [9]).

Глава 15

Задачи для самостоятельного решения

В последнем разделе собраны несложные задачи, решение которых поможет более уверенному овладению материалом курса. Предполагается, что из этих задач будут выбираться задания для приема зачета или экзамена по курсу.

1. Пусть η – случайная величина с функцией распределения $F(x)$. Зададим случайный процесс формулой $\xi(t) = \eta + t$. Найдите конечномерные распределения этого процесса.
2. Пусть ξ, η – случайные величины, η имеет симметричное распределение, причем $\mathbf{P}(\eta = 0) = 0$. Найдите вероятность того, что случайный процесс $X(t) = t^2 + t\eta + \xi$ имеет наверняка возрастающие при $t > 0$ траектории.
3. Покажите, что стохастически эквивалентные процессы всегда имеют одинаковые конечномерные распределения.
4. Докажите, что процесс, стохастически эквивалентный стохастически непрерывному процессу, сам является стохастически непрерывным.
5. Докажите, что, если случайный процесс стохастически непрерывен на компактном множестве, то он на этом множестве

равномерно стохастически непрерывен (предварительно дайте строгое определение равномерной непрерывности процесса!).

6. Пусть $\xi(t)$ – стохастически непрерывный стационарный случайный процесс и он имеет независимые приращения. Известно, что $\xi(0) = 0$, и $\xi(1/2) - \xi(1/3)$ имеет равномерное распределение на отрезке $[0,1]$. Найдите распределение $\xi(t)$.
7. Докажите положительную определенность функции двух переменных $K(t, s) = \min\{t, s\}$, $t, s > 0$.
8. Является ли положительно определенной функция

$$K(t, s) = t(1 - s), \quad t, s \in [0, 1]?$$

9. Для случайного процесса $\xi(t) = X \sin \omega t$, где ω – постоянная частота, $\mathbf{M}X = 1$, $\mathbf{D}X = 0,2$ найдите математическое ожидание и корреляционную функцию.
10. Пусть $\xi(t) = X^2 e^{-t^2/2}$, где X – нормально распределенная случайная величина с параметрами 2 и 0,01. Найдите математическое ожидание и корреляционную функцию этого случайного процесса.
11. Решите предыдущую задачу, если распределение X является равномерным, $\mathbf{M}X = 2$, $\mathbf{D}X = 1$.
12. Пусть U, V – некоррелированные случайные величины,

$$\mathbf{M}U = 0,5, \mathbf{M}V = 0,5; \mathbf{D}U = 1, \mathbf{D}V = 0,05.$$

Для случайного процесса $X(t) = Ut + Vt^2$ найдите математическое ожидание и корреляционную функцию.

13. Дан случайный процесс

$$\xi(t) = 2U \sin \omega t + 3Vt^2 + 5,$$

где U, V – случайные величины; $\mathbf{M}U = 1, \mathbf{M}V = 2, \mathbf{D}U = 0,1, \mathbf{D}V = 0,5, \rho(U, V) = -0,3$. Найдите его математическое ожидание и ковариационную функцию.

14. Случайный процесс $X(t)$ задан каноническим разложением

$$X(t) = t - 3 \cos t + U(t + \cos t) + V \cos 2t,$$

$\mathbf{D}U = 1$, $\mathbf{D}V = 2$. Найдите $\mathbf{M}X(t)$, $K(t, s)$, $\mathbf{D}X(t)$.

15. Каково математическое ожидание процесса Пуассона с параметром λ ? Найдите также его ковариационную функцию.

16. Случайные процессы заданы своими каноническими разложениями

$$X(t) = t + V \cos t + U \sin t, \quad Y(t) = 2t^2 - V \sin t + U \cos t,$$

где $\mathbf{D}U = \mathbf{D}V = \sigma^2$. Найдите математическое ожидание и ковариационную функцию суммы этих процессов $Z(t) = X(t) + Y(t)$, а также их взаимную ковариационную функцию

$$K_{X,Y}(t, s) = \mathbf{cov}(X(t), Y(s))$$

17. Дан случайный процесс $X(t) = x_1 t + x_2 \sin t$, где случайный вектор (x_1, x_2) имеет математическое ожидание $(1, -1)$ и ковариационную матрицу

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

Постройте каноническое разложение этого случайного процесса. Найдите $\mathbf{M}X(t)$, $K(t, s)$.

18. На плоскости движется случайная точка M так, что ее полярный угол φ является случайной функцией времени с ковариационной функцией

$$K(t, s) = a^2 \exp \{-b^2(t - s)^2\}.$$

Найдите дисперсию угловой скорости полярного радиус-вектора точки M .

19. Пусть $\xi(t)$ – гауссовский стационарный случайный процесс, $\mathbf{M}\xi(t) = 0$, и его ковариационная функция непрерывна. Зафиксируем некоторое действительное значение t . Найдите ковариационную функцию процесса

$$X(s) = \xi(t)\xi(t + s).$$

20. Задана матрица переходных вероятностей цепи Маркова

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix}.$$

Произведите наиболее полную из возможных здесь классификацию состояний.

21. Задана матрица переходных вероятностей цепи Маркова

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}.$$

Произведите наиболее полную классификацию ее состояний.

22. Во время случайного блуждания по целочисленным точкам прямой частица делает один шаг влево с вероятностью q , а вправо с вероятностью $p = 1 - q$ вне зависимости от предыстории своего движения. В получающейся при этом цепи Маркова вычислите $p_{00}^{(n)}$.

23. Пусть после проведения $n > 1$ испытаний Бернулли мы говорим, что система находится в состоянии E1, если последнее и предпоследнее испытания завершились успехами (УУ), E2, если успехом и неудачей (УН), E3, если НУ и E4, если НН. Найдите вероятности перехода между всеми возможными парами состояний за два шага. Является ли описанный процесс цепью Маркова?

24. Матрица перехода цепи Маркова из m состояний имеет вид

$$P = \begin{pmatrix} q & p & 0 & 0 & \dots & 0 \\ q & 0 & p & 0 & \dots & 0 \\ 0 & q & 0 & p & \dots & 0 \\ \dots & & & & \dots & \\ 0 & 0 & 0 & 0 & q & p \end{pmatrix}.$$

Найдите $\lim_{n \rightarrow \infty} P^n$.

25. Докажите, что в цепи Маркова для любого возвратного состояния можно указать множество состояний Γ , содержащее это состояние, такое, что для произвольных $i, k \in \Gamma$ справедливо

$$\lim_{n \rightarrow \infty} p_{ik}(n) = 1, \quad \lim_{n \rightarrow \infty} p_{ki}(n) = 1.$$

26. Докажите, что в конечной неразложимой цепи Маркова нет нулевых состояний.
27. Докажите, что в конечной цепи Маркова все состояния не могут одновременно быть невозвратными.
28. Пусть $X(t), t > 0$ – марковский процесс. Докажите, что последовательность $\xi_n = X(n), n \in \mathbf{N}$ образует цепь Маркова.
29. Будет ли процесс Пуассона непрерывен? Дифференцируем в среднем квадратичном?
30. Известны характеристики случайного процесса $X(t)$:

$$\mathbf{M}X(t) = 2t + 1; \quad K(t, s) = \exp\{-(t - s)^2\}.$$

Вычислите математическое ожидание и ковариационную функцию его производной.

31. Докажите, что случайный процесс

$$X(t) = e^{-at} \sin(\omega t + \varphi),$$

где a, ω – положительные постоянные, φ имеет равномерное распределение на $[0, 2\pi]$, дифференцируем при всех $t > 0$.

32. Случайный процесс $X(t)$ задан каноническим разложением

$$X(t) = 1 + t + Ut + Vt^2 + Wt^3,$$

где $\mathbf{D}U = 2, \mathbf{D}V = 1, \mathbf{D}W = 0, 1$. Определите характеристики производной процесса $\frac{dX}{dt}$.

33. Дан случайный процесс $X(t) = x_1 t + x_2 \sin t$, где случайный вектор (x_1, x_2) имеет математическое ожидание $(1, -1)$ и ковариационную матрицу

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

Вычислите характеристики случайного процесса

$$Y(t) = \int_0^t X(s) ds.$$

34. На вход интегрирующего устройства поступает случайная функция $X(t)$, $\mathbf{M}X(t) = 0, 2 \cos^2 \omega t$, его ковариационная функция $K(t, s) = 0, 4 \cos \omega t \cos \omega s$. Определите значение математического ожидания процесса на выходе интегратора. Как выглядит его ковариационная функция?
35. Дифференцируемый в среднем квадратичном случайный процесс $X(t)$ имеет ковариационную функцию $K(t, s)$,

$$Y(t) = X(t) + \frac{dX}{dt}.$$

Найдите ковариационную функцию $Y(t)$.

36. Пусть $X(t)$ – случайный процесс, $Y(t) = \int_0^t X(s) ds$. Определите, каким должен быть $X(t)$, чтобы $Y(t)$ был бы стационарным (хотя бы одно достаточное условие). Всегда ли интеграл стационарного случайного процесса стационарен?
37. В области G в начальный момент $t = 0$ имелось k частиц. Независимо друг от друга каждая из частиц за время t покидает область с вероятностью $\mu t + o(\Delta t)$. Новые частицы не появляются. Найдите математическое ожидание количества частиц, находящихся в области в момент времени t .
38. В области G имеются частицы, способные размножаться. Гибель или уменьшение числа частиц не отмечается. За промежуток времени t каждая частица независимо от остальных производит новую с вероятностью $\lambda t + o(\Delta t)$. Составьте прямую систему дифференциальных уравнений Колмогорова, определяющих этот процесс.
39. Процесс $\xi(t), t \in [a, b]$ стационарен. Докажите, что процесс

$$X(t) = \int_a^b B(t-s)\xi(s) ds,$$

где B – произвольная функция, для которой выписанный интеграл существует, также стационарен. Чему равна ковариационная функция этого процесса?

40. Стационарный случайный процесс имеет ковариационную функцию $K(t) = e^{-|t|} \cos \beta t$. Найдите его спектральную плотность.
41. Стационарный случайный процесс имеет спектральную плотность

$$f(\lambda) = \frac{\alpha}{\pi(\alpha^2 + \lambda^2)}.$$

Вычислите его ковариационную функцию.

42. Пусть ξ – случайная величина. Случайный процесс тождественно по t равен ξ . Можно ли говорить о наличии здесь некоторой спектральной меры и, если да, то как устроена эта спектральная мера?
43. Пусть ξ_1, ξ_2, \dots – независимые случайные величины, причем каждая из них имеет математическое ожидание, равное единице. Докажите, что

$$X_n = \prod_{k=1}^n \xi_k, \quad n \in \mathbf{N} -$$

мартингал.

44. Пусть $\xi(t)$ – мартингал, $M\xi^2(t) < \infty$. Докажите, что он имеет некоррелированные приращения.
45. Пусть $X(t)$ – пуассоновский процесс с параметром λ . Докажите, что $Y(t) = \exp\{X(t) - at\}$ есть субмартингал при $a < \lambda t$ и супермартингал при $a > \lambda t$.
46. Среднее число выбросов гауссовского стационарного процесса $\xi(t)$ за уровень $a = M\xi(t)$ в единицу времени равно 0,01. Дисперсия процесса равна 64. Найдите дисперсию производной процесса.

Библиографический СПИСОК

- [1] Robinson A. Non-standard analysis. – Princeton University Press, 1996, – 308 P.
- [2] Zadeh L. A. Fuzzy sets. // Information and Control. – 1965. – V. 8, № 3. – P. 338 – 353.
- [3] Дронов С.В. Элементы прикладной альтернативной теории множеств. LAP Lambert Academic Publishing, Saarbrucken, Deutschland, 2012, – 183 P.
- [4] Боровков А.А. Теория вероятностей. Учебное пособие для вузов.– Либроком, 2009.– 656 с.
- [5] Семаков С.Л. Выбросы случайных процессов: учебное пособие. – М.: МАДИ (ГТУ), 2004. – 70 с.
- [6] Гнеденко Б.В. Курс теории вероятностей. 8-е изд., испр. и доп.– М.: Едиториал УРСС, 2005.– 448 с.
- [7] Ширяев А.Н. Вероятность. – МЦНМО, 2007. – 968 с.
- [8] Лаврентьев М.А., Шабат Б.В. Методы теории функций комплексного переменного. М.: Наука, 1973. – 736 с.
- [9] Лифшиц Е. М., Питаевский Л. П. Физическая кинетика.– М.: Наука, 1979. – 528 с.
- [10] Баскаков С. И. Радиотехнические цепи и сигналы. Изд. четвертое. URSS, 2022. – 528 с.

- [11] Булинский А.В., Ширяев А.Н. Теория случайных процессов. – М.: Физматлит, 2005. – 400 с.
- [12] Вентцель А.Д. Курс теории случайных процессов. – М.: Наука, 1996. – 400 с.
- [13] Вентцель Е. С., Овчаров Л. А. Теория случайных процессов и её инженерные приложения. – М.: Наука, 1991. – 384 с.
- [14] Гихман И.И., Скороход А.В. Введение в теорию случайных процессов.– М.: Наука, 1977. – 570 с.
- [15] Гихман И.И., Скороход А.В. Теория случайных процессов в 3 томах. – М: Наука, 1971 – 1975.
- [16] Карлин С. Основы теории случайных процессов. – ЕЕ Медиа, 2013. – 537 с.
- [17] Миллер Б.М., Панков А.Р. Теория случайных процессов в примерах и задачах. – М.: Физматлит, 2007. – 320 с.
- [18] Петунин Ю.И. Приложение теории случайных процессов в биологии и медицине. – Киев: Наукова думка, 1981. – 320 с.
- [19] Портенко Н.И., Скороход А.В., Шуренков В.М. Марковские процессы.// Итоги науки и техники. Серия Современные проблемы математики. Фундаментальные направления. Т. 46. Теория вероятностей-4. – М.: ВИНТИ, 1989. – 248 с.
- [20] Розанов Ю.А. Введение в теорию случайных процессов. М.: Наука, 1982. – 130 с.